

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2006, 13:07:33 ; Search time 233 Seconds  
(without alignments)  
2240.734 Million cell updates/sec

Title: US-10-659-055-3  
Perfect score: 4020  
Sequence: 1 ADPGSHHHHRSKTYTLT.....AHQHIYTHSHFIKQCFSLP 740

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3939	98.0	766	1 DPP4_HUMAN	P27487 h dipeptidyl
2	3939	98.0	766	2 Q53TN1_HUMAN	Q53tn1 homo sapien
3	3910	97.3	765	2 Q5R7G7_PONPY	Q5r7g7 pongo pygma
4	3585	89.2	765	1 DPP4_BOVIN	P81425 b dipeptidyl
5	3560	88.6	765	1 DPP4_FELCA	Q9n217 f dipeptidyl
6	3548	88.3	766	1 DPP4_PIG	P22411 s dipeptidyl
7	3408.5	84.8	767	1 DPP4_RAT	P14740 r dipeptidyl
8	3390	84.3	760	1 DPP4_MOUSE	P28843 m dipeptidyl
9	2520.5	62.7	751	2 Q75S83_AGKHB	Q75s83 agkistrodon
10	2519.5	62.7	751	2 Q75S82_AGKHB	Q75s82 agkistrodon
11	2519	62.7	759	2 Q5Z1B1_CHICK	Q5zi81 gallus gall
12	2348	58.4	737	2 Q641D6_XENLA	Q641d6 xenopus lae
13	2331.5	58.0	748	2 P70092_XENLA	P70092 xenopus lae
14	2175	54.1	760	2 Q53TP5_HUMAN	Q53tp5 homo sapien
15	2168	53.9	760	1 SEPR_HUMAN	Q12884 homo sapien
16	2160.5	53.7	761	2 Q8R452_RAT	Q8r492 rattus norv
17	2158.5	53.7	761	1 SEPR_MOUSE	P97321 mus musculu
18	2115	52.6	755	2 Q91651_XENLA	Q91651 xenopus lae
19	1992.5	49.6	860	2 Q4RQJ4_TETNG	Q4rqj4 tetraodon n
20	1819.5	45.3	704	2 Q4S309_TETNG	Q4s309 tetraodon n
21	1222	30.4	796	1 DPP10_HUMAN	Q8n608 homo sapien
22	1210	30.1	796	1 DPP10_RAT	Q8g629 rattus norv
23	1202.5	29.9	796	2 Q6INB7_XENLA	Q6inb7 xenopus lae
24	1183	29.4	797	1 DPP10_MOUSE	Q6nxt7 mus musculu
25	1129	28.1	865	1 DPP6_HUMAN	P42658 homo sapien
26	1124	28.0	731	2 Q7PSF9_ANOGA	Q7psf9 anopheles g
27	1121	27.9	803	1 DPP6_PANTR	Q7is50 pan troglod
28	1118	27.8	803	2 Q9QV78_9MURI	Q9qv78 rattus sp.
29	1116	27.8	859	1 DPP6_RAT	P46101 rattus norv
30	1112.5	27.7	745	2 Q9VMM2_DROME	Q9vmm2 drosophila
31	1106	27.5	863	1 DPP6_BOVIN	P42659 bos taurus

32	1089	27.1	801	2 Q5U4C2_MOUSE	Q5u4c2 mus musculu
33	1089	27.1	803	2 Q80VM5_MOUSE	Q80vm5 mus musculu
34	1089	27.1	804	1 DPP6_MOUSE	Q92218 mus musculu
35	1088.5	27.1	439	2 Q6PG25_XENLA	Q6pg25 xenopus lae
36	1069.5	26.6	815	2 Q4RK63_TETNG	Q4rk63 tetraodon n
37	1059	26.3	724	2 Q5IS95_9PRIM	Q5is95 saimiri bol
38	1051	26.1	604	2 Q53S84_HUMAN	Q53s84 homo sapien
39	1045	26.0	801	2 Q4RK66_TETNG	Q4rk66 tetraodon n
40	1038	25.8	802	2 Q9VUH1_DROME	Q9vuh1 drosophila
41	1025.5	25.5	935	2 Q3VMB4_DROME	Q3vmb4 drosophila
42	998	24.8	775	2 Q5J6J3_TIRUR	Q5j6j3 trichophyco
43	996	24.8	842	2 Q4RVF8_TETNG	Q4rvf8 tetraodon n
44	994	24.7	874	2 Q7SI80_EMENI	Q7si80 emericella
45	986	24.5	880	2 Q5B934_EMENI	Q5b934 aspergillus

#### ALIGNMENTS

RESULT 1  
DPP4\_HUMAN STANDARD; PRT; 766 AA.  
AC P27487;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP  
DE IV) (T-cell activation antigen CD26) (TP103) (Adenosine deaminase  
DE complexing protein 2) (ADAPF) [Contains: Dipeptidyl peptidase 4  
DE membrane form (Dipeptidyl peptidase IV membrane form); Dipeptidyl  
DE peptidase 4 soluble form (Dipeptidyl peptidase IV soluble form)].  
GN Name=DPP4; Synonyms=ADCP2, CD26;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Liver;  
RX MEDLINE=92329551; PubMed=1352704; DOI=10.1016/0167-4781(92)90036-Y;  
RA Misumi Y., Hayashi Y., Arakawa F., Ikehara Y.;  
RT "Molecular cloning and sequence analysis of human dipeptidyl peptidase  
RT IV, a serine proteinase on the cell surface.";  
RL Blochim. Biophys. Acta 1131:333-336(1992).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Placenta;  
RX MEDLINE=95012454; PubMed=7927537;  
RA Abbott C.A., Baker E., Sutherland G.R., McCaughan G.W.;  
RT "Genomic organization, exact localization, and tissue expression of  
RT the human CD26 (dipeptidyl peptidase IV) gene.";  
RL Immunogenetics 40:331-338(1994).  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Periphereal blood;  
RX MEDLINE=92325476; PubMed=1352530;  
RA Tanaka T., Camerini D., Seed B., Torimoto Y., Dang N.H., Kameoka J.,  
RA Dahlberg H.N., Schlossman S.F., Morimoto C.;  
RT "Cloning and functional expression of the T cell activation antigen  
RT CD26.";  
RL J. Immunol. 149:481-486(1992).  
RN [4]  
RP ERRATUM.  
RX MEDLINE=93171637; PubMed=8094732;  
RA Tanaka T.;  
RN [5]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
RC TISSUE=Prostate, and Uterus;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altechul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Boeak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.C., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green B.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywicki M.I., Skalek U., Smalius D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [6]  
RP NUCLEOTIDE SEQUENCE OF 1-551.  
RC TISSUE=Colon;  
RX MEDLINE=92165847; PubMed=1347043;  
RA Darnoul D., Lacasa M., Baricault L., Marguet D., Sapin C., Trotot P.,  
RA Barbat A.;  
RT "Dipeptidyl peptidase IV (CD 26) gene expression in enterocyte-like  
RT colon cancer cell lines HT-29 and Caco-2. Cloning of the complete  
RT human coding sequence and changes of dipeptidyl peptidase IV mRNA  
RT levels during cell differentiation.";  
RL J. Biol. Chem. 267:4824-4833(1992).  
RN [7]  
RP NUCLEOTIDE SEQUENCE OF 545-766.  
RC TISSUE=Colon;  
RX MEDLINE=91024044; PubMed=1977364;  
RA Darnoul D., Lacasa M., Chantret I., Swallow D., Trugnan G.;  
RT "Isolation of a cDNA probe for the human intestinal  
RT dipeptidylpeptidase IV and assignment of the gene locus DPP4 to  
RT chromosome 2.";  
RL Ann. Hum. Genet. 54:191-197(1990).  
RN [8]  
RP NUCLEOTIDE SEQUENCE OF 1-31.  
RX MEDLINE=96067599; PubMed=7487939;  
RA Boehm S.K., Gum J.R. Jr., Erickson R.H., Hicks J.W., Kim Y.S.;  
RT "Human dipeptidyl peptidase IV gene promoter: tissue-specific  
RT regulation from a TATA-less GC-rich sequence characteristic of a  
RT housekeeping gene promoter.";  
RL Biochem. J. 311:835-843(1995).  
RN [9]  
RP PROTEIN SEQUENCE OF 1-22, AND TISSUE SPECIFICITY.  
RX PubMed=1677636;  
RA Gorvel J.P., Ferrero A., Chambrud L., Rigal A., Bonicel J.,  
RA Maroux S.;  
RT "Expression of sucrase-isomaltase and dipeptidylpeptidase IV in human  
RT small intestine and colon.";  
RL Gastroenterology 101:618-625(1991).  
RN [10]  
RP PARTIAL PROTEIN SEQUENCE.  
RC TISSUE=Kidney;  
RX MEDLINE=93210468; PubMed=8096237; DOI=10.1084/jem.177.4.1135;  
RA Morrison M.E., Vijayasaradhi S., Engelstein D., Albino A.P.,  
RA Houghton A.N.;  
RT "A marker for neoplastic progression of human melanocytes is a cell  
RT surface ectopeptidase.";  
RL J. Exp. Med. 177:1135-1143(1993).  
RN [11]  
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 38-766 IN COMPLEX WITH  
RP INHIBITOR, AND HOMODIMERIZATION.  
RX PubMed=12832764; DOI=10.1107/S0907444903010059;  
RA Oefner C., D'Arcy A., Mac Sweeney A., Pirau S., Gardiner R.,  
RA Dale G.E.;  
RT "High-resolution structure of human apo dipeptidyl peptidase IV/CD26  
RT and its complex with 1-[(2-[(5-iodopyridin-2-yl)amino]-ethyl)amino]-  
RT acetyl]-2-cyano-(S)-pyrrolidine.";  
RL Acta Crystallogr. D 59:1206-1212(2003).  
RN [12]  
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-771, HOMODIMERIZATION, AND  
RP N-GLYCOSYLATION SITES.  
RX PubMed=12846248; DOI=10.1016/S0006-291X(03)00258-4;  
RA Hiragatsu H., Kyono K., Higashiyama Y., Fukushima C., Shima H.,  
RA Sugiyama S., Inaka K., Yamamoto A., Shimizu R.;  
RT "The structure and function of human dipeptidyl peptidase IV,  
RT possessing a unique eight-bladed beta-propeller fold.";  
RL Biochem. Biophys. Res. Commun. 302:849-854(2003).  
RN [13]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 39-766, HOMODIMERIZATION, AND  
RP N-GLYCOSYLATION SITES.  
RX PubMed=12483204; DOI=10.1038/nab882;  
RA Rasmussen H.B., Branner S., Wiberg P.C., Wagtmann N.;  
RT "Crystal structure of human dipeptidyl peptidase IV/CD26 in complex  
RT with a substrate analog.";  
RL Nat. Struct. Biol. 10:19-25(2003).  
RN [14]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 39-766, HOMODIMERIZATION, AND  
RP N-GLYCOSYLATION SITES.  
RX PubMed=12906826; DOI=10.1016/S0969-2126(03)00160-6;  
RA Thoma R., Loeffler B., Stihle M., Huber W., Ruf A., Hennig M.;  
RT "Structural basis of proline-specific exopeptidase activity as  
RT observed in human dipeptidyl peptidase-IV.";  
RL Structure 11:947-959(2003).  
CC -|- FUNCTION: Removes N-terminal dipeptides sequentially from the  
CC polypeptides having unsubstituted N-termini provided that the  
CC penultimate residue is proline. Plays a role in T cell activation.  
CC -|- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-  
CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided  
CC Zaa is neither Pro nor hydroxyproline.  
CC -|- SUBUNIT: Homodimer or heterodimer with seprase (PAP).  
CC -|- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a  
CC soluble form.  
CC -|- TISSUE SPECIFICITY: Expressed in the poorly differentiated crypt  
CC cells of the small intestine as well as in the mature villous  
CC cells. Expressed at very low levels in the colon.  
CC -|- PTM: The soluble form (SDPP) derives from the membrane form (MDPP)  
CC by proteolytic processing.  
CC -|- SIMILARITY: Belongs to the peptidase 98B family. DPPIV subfamily.  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
CC EMBL; U13735; AAB60646.1; -; Genomic DNA.  
CC EMBL; U13710; AAB60646.1; JOINED; Genomic DNA.  
CC EMBL; U13711; AAB60646.1; JOINED; Genomic DNA.  
CC EMBL; U13712; AAB60646.1; JOINED; Genomic DNA.  
CC EMBL; U13713; AAB60646.1; JOINED; Genomic DNA.  
CC EMBL; U13714; AAB60646.1; JOINED; Genomic DNA.  
CC EMBL; U13715; AAB60646.1; JOINED; Genomic DNA.  
CC EMBL; U13716; AAB60646.1; JOINED; Genomic DNA.  
CC EMBL; U13717; AAB60646.1; JOINED; Genomic DNA.  
CC EMBL; U13718; AAB60646.1; JOINED; Genomic DNA.  
CC EMBL; U13719; AAB60646.1; JOINED; Genomic DNA.  
CC EMBL; U13720; AAB60646.1; JOINED; Genomic DNA.  
CC EMBL; U13721; AAB60646.1; JOINED; Genomic DNA.  
CC EMBL; U13722; AAB60646.1; JOINED; Genomic DNA.  
CC EMBL; U13723; AAB60646.1; JOINED; Genomic DNA.  
CC EMBL; U13724; AAB60646.1; JOINED; Genomic DNA.  
CC EMBL; U13725; AAB60646.1; JOINED; Genomic DNA.  
CC EMBL; U13726; AAB60646.1; JOINED; Genomic DNA.  
CC EMBL; U13727; AAB60646.1; JOINED; Genomic DNA.  
CC EMBL; U13728; AAB60646.1; JOINED; Genomic DNA.  
CC EMBL; U13729; AAB60646.1; JOINED; Genomic DNA.  
CC EMBL; U13730; AAB60646.1; JOINED; Genomic DNA.  
CC EMBL; U13731; AAB60646.1; JOINED; Genomic DNA.  
CC EMBL; U13732; AAB60646.1; JOINED; Genomic DNA.  
CC EMBL; U13733; AAB60646.1; JOINED; Genomic DNA.  
CC EMBL; U13734; AAB60646.1; JOINED; Genomic DNA.

DR EMBL; M74777; AAA51943.1.; -; mRNA.  
DR EMBL; BC013329; AAH13329.2.; -; mRNA.

Query Match 98.0%; Score 3939; DB 1; Length 766;  
Best Local Similarity 100.0%; Pred. No. 2e-259;  
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 SRKTYTLDYLNKTYRLKLYSLRWISDHEYLKQENNLVFNABYGNSSVPLENSTDEF 72  
Db |||||  
Qy 39 SRKTYTLDYLNKTYRLKLYSLRWISDHEYLKQENNLVFNABYGNSSVPLENSTDEF 98  
Db |||||

Qy 73 GHSINDYSISPDGQFILLEYNVYKQWHSYTSASDIYDLNKRQLITEERIPNNTQWTVS 132  
Db |||||

Qy 99 GHSINDYSISPDGQFILLEYNVYKQWHSYTSASDIYDLNKRQLITEERIPNNTQWTVS 158  
Db |||||

Qy 133 PVGHKLAYVWNNDIYVYKIEPNLPSYRITWTGKEDIYNGITDWTYBEVFSAYGALWSP 192  
Db |||||

Qy 159 PVGHKLAYVWNNDIYVYKIEPNLPSYRITWTGKEDIYNGITDWTYBEVFSAYGALWSP 218  
Db |||||

Qy 193 NGTFLAYAQFNDTEVPLIEYSFSDLSIQPKTVRVPYKAGAVNPTVKFFVNTDLSLS 252  
Db |||||

Qy 219 NGTFLAYAQFNDTEVPLIEYSFSDLSIQPKTVRVPYKAGAVNPTVKFFVNTDLSLS 278  
Db |||||

Qy 253 VTNATSIQITAPASMLIGDHYLGDVWTATQERISLOWLRRIQNYSVMDICDYDESSGRWN 312  
Db |||||

Qy 279 VTNATSIQITAPASMLIGDHYLGDVWTATQERISLOWLRRIQNYSVMDICDYDESSGRWN 338  
Db |||||

Qy 313 CLVARQHIEMSTTGWGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDKCTFIT 372  
Db |||||

Qy 339 CLVARQHIEMSTTGWGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDKCTFIT 398  
Db |||||

Qy 373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQVYS 432  
Db |||||

Qy 399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQVYS 458  
Db |||||

Qy 433 VFSKEAKYQLRCGSGPLGYTLHSSVNDKGLRVLEDNSALDKQLQNVQWPSKKLDFII 492  
Db |||||

Qy 459 VFSKEAKYQLRCGSGPLGYTLHSSVNDKGLRVLEDNSALDKQLQNVQWPSKKLDFII 518  
Db |||||

Qy 493 LNETKFWYQMLPPHFDKSKYKPLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASF 552  
Db |||||

Qy 519 LNETKFWYQMLPPHFDKSKYKPLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASF 578  
Db |||||

Qy 553 DGRSGYQGDKIMHAINRRLGTPEVDQIEAARQFSKMGFVNDKRIAIWGSYGYVTSM 612  
Db |||||

Qy 579 DGRSGYQGDKIMHAINRRLGTPEVDQIEAARQFSKMGFVNDKRIAIWGSYGYVTSM 638  
Db |||||

Qy 613 VLGSYGVPKCGIAVAPVSRWEYDSTYTERYMGGLPTPEDNLDHYRNSVTMSRAENPKQV 672  
Db |||||

Qy 639 VLGSYGVPKCGIAVAPVSRWEYDSTYTERYMGGLPTPEDNLDHYRNSVTMSRAENPKQV 698  
Db |||||

Qy 673 EYLLIHGTADNVHFQSAQISKALVDGVDFQAMWYTDDEHGTASSTAHQHIYTHMSHF 732  
Db |||||

Qy 699 EYLLIHGTADNVHFQSAQISKALVDGVDFQAMWYTDDEHGTASSTAHQHIYTHMSHF 758  
Db |||||

Qy 733 IKQCFSLP 740  
Db |||||

Qy 759 IKQCFSLP 766  
Db |||||

RESULT 2  
Q53TN1 HUMAN  
ID Q53TN1 HUMAN PRELIMINARY; PRT; 766 AA.  
AC Q53TN1;  
DT 13-SEP-2005 (TrEMBLrel. 31, Created)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Hypothetical protein DPP4.  
GN Name=DPP4;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.

NCBI\_TaxID=9606;  
[1] NUCLEOTIDE SEQUENCE.  
RP Ozanich A., Stoneking T., Hawkins M., Sapetti L.;  
RA "The sequence of Homo sapiens BAC clone RP11-178A14.";  
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
[2] NUCLEOTIDE SEQUENCE.  
RP Waterston R.H.;  
RA Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
[3] NUCLEOTIDE SEQUENCE.  
RP Waterston R.;  
RA Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
[4] NUCLEOTIDE SEQUENCE.  
RP Wilson R.K.;  
RA Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AC008063; AA931179.1.; -; Genomic\_DNA.  
DR SMR; Q53TN1; 39-766.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0003824; F:catalytic activity; IEA.  
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.  
DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR002471; Pept\_S9 AS.  
DR InterPro; IPR001375; Peptidase\_S9.  
DR InterPro; IPR002469; Peptidase\_S9B.  
DR InterPro; IPR000379; Ser\_eatra.  
DR Pfam; PF00930; DPPIV\_N; 1.  
DR Pfam; PF00326; Peptidase\_S9; 1.  
DR PROSITE; PS00708; PRO\_ENDOPSP SER; 1.  
KW Hydrolase; Hypothetical protein; Transmembrane.  
SQ SEQUENCE 766 AA; 88279 MW; 5FB4A2C6662D6117 CRC64;

Query Match 98.0%; Score 3939; DB 2; Length 766;  
Best Local Similarity 100.0%; Pred. No. 2e-259;  
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 SRKTYTLDYLNKTYRLKLYSLRWISDHEYLKQENNLVFNABYGNSSVPLENSTDEF 72  
Db |||||

Qy 39 SRKTYTLDYLNKTYRLKLYSLRWISDHEYLKQENNLVFNABYGNSSVPLENSTDEF 98  
Db |||||

Qy 73 GHSINDYSISPDGQFILLEYNVYKQWHSYTSASDIYDLNKRQLITEERIPNNTQWTVS 132  
Db |||||

Qy 99 GHSINDYSISPDGQFILLEYNVYKQWHSYTSASDIYDLNKRQLITEERIPNNTQWTVS 158  
Db |||||

Qy 133 PVGHKLAYVWNNDIYVYKIEPNLPSYRITWTGKEDIYNGITDWTYBEVFSAYGALWSP 192  
Db |||||

Qy 159 PVGHKLAYVWNNDIYVYKIEPNLPSYRITWTGKEDIYNGITDWTYBEVFSAYGALWSP 218  
Db |||||

Qy 193 NGTFLAYAQFNDTEVPLIEYSFSDLSIQPKTVRVPYKAGAVNPTVKFFVNTDLSLS 252  
Db |||||

Qy 219 NGTFLAYAQFNDTEVPLIEYSFSDLSIQPKTVRVPYKAGAVNPTVKFFVNTDLSLS 278  
Db |||||

Qy 253 VTNATSIQITAPASMLIGDHYLGDVWTATQERISLOWLRRIQNYSVMDICDYDESSGRWN 312  
Db |||||

Qy 279 VTNATSIQITAPASMLIGDHYLGDVWTATQERISLOWLRRIQNYSVMDICDYDESSGRWN 338  
Db |||||

Qy 313 CLVARQHIEMSTTGWGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDKCTFIT 372  
Db |||||

Qy 339 CLVARQHIEMSTTGWGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDKCTFIT 398  
Db |||||

Qy 373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQVYS 432  
Db |||||

Qy 399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQVYS 458  
Db |||||

Qy 433 VFSKEAKYQLRCGSGPLGYTLHSSVNDKGLRVLEDNSALDKQLQNVQWPSKKLDFII 492  
Db |||||

Qy 459 VFSKEAKYQLRCGSGPLGYTLHSSVNDKGLRVLEDNSALDKQLQNVQWPSKKLDFII 518  
Db |||||

Qy 493 LNETKFWYQMLPPHFDKSKYKPLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASF 552  
Db |||||

```
Db 519 LNETKFWYQMLPPHFDKSKYPPLLLDVYAGPCSKADTVPRLNWATYLASTENIIVASF 578
Qy 553 DGRSGYGQDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKIAIWGWSYGGYVTSM 612
Db 579 DGRSGYGQDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKIAIWGWSYGGYVTSM 638
Qy 613 VLGGSGGVKFCGIAVAPVSRWEYDVSVYTERVMGLPTPEDNLDHYRNSVMSRAENFKQV 672
Db 639 VLGGSGGVKFCGIAVAPVSRWEYDVSVYTERVMGLPTPEDNLDHYRNSVMSRAENFKQV 698
Qy 673 EYLLHGTADDNVHFQQAQISKALVDVGVDFOAMWYTDDEHGIIASSTAHOHYTHMSHF 732
Db 699 EYLLHGTADDNVHFQQAQISKALVDVGVDFOAMWYTDDEHGIIASSTAHOHYTHMSHF 758
Qy 733 IKQCFSLP 740
Db 759 IKQCFSLP 765

RESULT 3
QSR7G7 PONPY PRELIMINARY; PRT; 765 AA.
ID QSR7G7;
AC QSR7G7;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein DKFZp469P1419;
GN Name=DKFZp469P1419;
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC The German CNA Consortium;
RA Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,
RA Mewes H.W., Well B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBSJ databases.
DR EMBL; CR860150; CAH92293.1; -; mRNA.
DR SMR; QSR7G7; 38-765.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004177; P:aminopeptidase activity; IEA.
DR GO; GO:0004274; P:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0016787; P:hydrolase activity; IEA.
DR GO; GO:0004287; P:prolyl oligopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002471; Pept S9 AS.
DR InterPro; IPR001375; Peptidase S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00930; DDPV N; 1.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW Hydrolase, Hypothetical protein, Transmembrane.
SQ SEQUENCE 765 AA; 88038 MW; 6F81ECD98C3D397 CRC64;

Query Match 97.3%; Score 3910; DB 2; Length 765;
Best Local Similarity 99.2%; Pred. No. 1.9e-257;
Matches 722; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 13 SRKTYTLDTLQNTYRLKLYSLRWISDHEYLKQENNLVFNAYGNSVFLNSTFDF 72
Db 38 SRKTYTLDTLQNTYRLKLYSLRWISDHEYLKQENNLVFNAYGNSVFLNSTFDF 97
Qy 73 GHSINDYSTSPGQFTLLSYVYKWRHSYTSYDIYDLNKKQLITEERIPNNTQWTS 132
Db 98 GHSINDYSTSPGQFTLLSYVYKWRHSYTSYDIYDLNKKQLITEERIPNNTQWTS 157
Qy 133 PVGHKLAYWNNDIYVKIEPNLPSRITWTGKEDIYNGITDWTYBEEVFSAYSLWSP 192
Db 158 PVGHKLAYWNNDIYVKIEPNLPSRITWTGKEDIYNGITDWTYBEEVFSAYSLWSP 217
```

```
Qy 193 NGTFLAYAQPNDDTEVPLEYSFYSDESIQQYKTVRVVPYKAGAVNPTVKFFVNTDSLSS 252
Db 218 NGTFLAYAQPNDDTEVPLEYSFYSDESIQQYKTVRVVPYKAGAVNPTVKFFVNTDSLSS 277
Qy 253 VTNATSIQITAPASNLIGDHYLDCVWTATQBRISLQWLRRIQNSVMDICDYDESSGRWN 312
Db 278 VTNATSIQITAPASNLIGDHYLDCVWTATQBRISLQWLRRIQNSVMDICDYDESSGRWN 337
Qy 313 CLVARQHIEMTGTWVGFRFSEPHFTLDGNSFYKIIISNEGYRHCYFQIDKKDCTFTT 372
Db 338 CLVARQHIEMTGTWVGFRFSEPHFTSDGNSFYKIIISNEGYRHCYFQIDKKDCTFTT 397
Qy 373 KGTWEVIGIEALTSDLYYISNEYKMGPGGRNLYKIQLSDTYKVTCLSCELNPERCQYYS 432
Db 398 KGTWEVIGIEALTSDLYYISNEYKMGPGGRNLYKIQLSDTYKVTCLSCELNPERCQYYS 457
Qy 433 VFSKEAKYQIQRCSGPGPLTYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFTI 492
Db 458 VFSKEAKYQIQRCSGPGPLTYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFTI 517
Qy 493 LNETKFWYQMLPPHFDKSKYPPLLLDVYAGPCSKADTVPRLNWATYLASTENIIVASF 552
Db 518 LNETKFWYQMLPPHFDKSKYPPLLLDVYAGPCSKADTVPRLNWATYLASTENIIVASF 577
Qy 553 DGRSGYGQDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKIAIWGWSYGGYVTSM 612
Db 578 DGRSGYGQDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKIAIWGWSYGGYVTSM 637
Qy 613 VLGGSGGVKFCGIAVAPVSRWEYDVSVYTERVMGLPTPEDNLDHYRNSVMSRAENFKQV 672
Db 638 VLGGSGGVKFCGIAVAPVSRWEYDVSVYTERVMGLPTPEDNLDHYRNSVMSRAENFKQV 697
Qy 673 EYLLHGTADDNVHFQQAQISKALVDVGVDFOAMWYTDDEHGIIASSTAHOHYTHMSHF 732
Db 698 EYLLHGTADDNVHFQQAQISKALVDVGVDFOAMWYTDDEHGIIASSTAHOHYTHMSHF 757
Qy 733 IKQCFSLP 740
Db 758 IKQCFSLP 765

RESULT 4
DPP4_BOVIN
ID DPP4_BOVIN STANDARD; PRT; 765 AA.
AC P81425; Q8WMSG;
DT 15-DEC-1998 (Rel. 37, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP
DE IV) (T-cell activation antigen CD26) (Adenosine deaminase complexing
DE protein) (ADCP-I) (Activation molecule 3) (ACT3) (WC10) [Contains:
DE Dipeptidyl peptidase 4 membrane form (Dipeptidyl peptidase IV membrane
DE form); Dipeptidyl peptidase 4 soluble form (Dipeptidyl peptidase IV
DE soluble form)].
DE Name=DPP4; Synonyms=CD26;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Lymphocyte;
RX MEDLINE=22067734; PubMed=12073152; DOI=10.1007/s00251-002-0456-6;
RA Lee S.-U., Park Y.-H., Davis W.C., Hamilton M.J., Naessens J.,
RA Bohach G.A.;
RT "Molecular characterization of bovine CD26 upregulated by a
RT staphylococcal superantigen.";
RL Immunogenetics 54:216-220 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 1-17.
RC TISSUE=Thymus;
```



RX MEDLINE=22021197; PubMed=11981836;  
 RX DOI=10.1002/1521-4141(200205)32:5<1472::AID-IMMU1472>3.0.CO;2-Q;  
 RA Gliddon D.R., Howard C.J.;  
 RT "CD26 is expressed on a restricted subpopulation of dendritic cells in  
 RT vivo";  
 RL Eur. J. Immunol. 32:1472-1481(2002).  
 RN [3]  
 RP PROTEIN SEQUENCE OF 1-24.  
 RC TISSUE=T-cell;  
 RX MEDLINE=21482004; PubMed=11598101;  
 RX DOI=10.1128/JAI.69.11.7190-7193.2001;  
 RA Lee S.-U., Ferens W., Davis W.C., Hamilton M.J., Park Y.-H., Fox L.K.,  
 RA Naessens J., Bohach G.A.;  
 RT "Identity of activation molecule 3 on superantigen-stimulated bovine  
 RT cells is CD26.";  
 RL Infect. Immun. 69:7190-7193(2001).  
 RN [4]  
 RP PROTEIN SEQUENCE OF 537-546.  
 RC TISSUE=Kidney;  
 RX MEDLINE=98293306; PubMed=9629661; DOI=10.1016/S0305-0491(97)00327-1;  
 RA Ben-Shoshan I., Parola A.H.;  
 RT "The CP-I subunit of adenosine deaminase complexing protein from calf  
 RT kidney is identical to human, mouse, and rat dipeptidyl peptidase  
 RT IV.";  
 RL Comp. Biochem. Physiol. 119B:289-292(1998).  
 CC -I- FUNCTION: Removes N-terminal dipeptides sequentially from  
 CC polypeptides having unsubstituted N-termini provided that the  
 CC penultimate residue is proline. Binds and regulates the activity  
 CC of ADA.  
 CC -I- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-  
 CC Zaa, from a polypeptide, preferentially when Yaa is Pro, provided  
 CC Zaa is neither Pro nor hydroxyproline.  
 CC -I- SUBUNIT: Homodimer.  
 CC -I- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a  
 CC soluble form (By similarity).  
 CC -I- TISSUE SPECIFICITY: Intestinal epithelium, dendritic cells and  
 CC several immune system tissues.  
 CC -I- PTM: The soluble form (SDPP) derives from the membrane form (MDPP)  
 CC by proteolytic processing (By similarity).  
 CC -I- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.  
 CC  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC  
 CC -----  
 DR EMBL; AF461806; AAL67836.1; -; mRNA.  
 DR EMBL; AY056834; AAL23628.1; -; mRNA.  
 DR HSSP; P27487; 1PFO.  
 DR SMR; P81425; 38-764.  
 DR MEROPS; S09.003; -.  
 DR InterPro; IPR002471; Pept\_S9\_AS.  
 DR InterPro; IPR001375; Peptidase S9.  
 DR InterPro; IPR002469; Peptidase S9B.  
 DR InterPro; IPR000379; Ser estrs.  
 DR Pfam; PF00930; DPPIV N; 1.  
 DR Pfam; PF00326; Peptidase S9; 1.  
 DR PROSITE; PS00708; PRO-ENDOPEP\_SER; 1.  
 DR AminoPeptidase; Direct protein sequencing; Glycoprotein; Hydrolase;  
 KW Protease; Serine protease; Signal-anchor; Transmembrane.  
 FT CHAIN 1 765  
 FT FT Dipeptidyl peptidase 4 membrane form.  
 FT FT Dipeptidyl peptidase 4 soluble form (By  
 FT FT similarity).  
 FT FT Cytoplasmic (Potential).  
 FT FT Signal-anchor for type II membrane  
 FT FT protein (Potential).  
 FT FT Extracellular (Potential).  
 FT FT Charge relay system (By similarity).  
 FT FT ACT\_SITE 629 629  
 FT FT Charge relay system (By similarity).  
 FT FT ACT\_SITE 707 707  
 FT FT Charge relay system (By similarity).  
 FT FT ACT\_SITE 739 739  
 FT FT Charge relay system (By similarity).  
 FT FT CARBOHYD 84 84  
 FT FT N-linked (GlcNAc...) (By similarity).  
 FT FT CARBOHYD 91 91  
 FT FT N-linked (GlcNAc...) (By similarity).

FT	CARBOHYD	149	149	N-linked (GlcNAc...)	(Potential).
FT	CARBOHYD	218	218	N-linked (GlcNAc...)	(By similarity).
FT	CARBOHYD	228	228	N-linked (GlcNAc...)	(By similarity).
FT	CARBOHYD	271	271	N-linked (GlcNAc...)	(Potential).
FT	CARBOHYD	280	280	N-linked (GlcNAc...)	(By similarity).
FT	CARBOHYD	320	320	N-linked (GlcNAc...)	(By similarity).
FT	CARBOHYD	392	392	N-linked (GlcNAc...)	(Potential).
FT	CARBOHYD	495	495	N-linked (GlcNAc...)	(Potential).
FT	CARBOHYD	684	684	N-linked (GlcNAc...)	(By similarity).
FT	DISULFID	384	393	By similarity.	
FT	DISULFID	443	446	By similarity.	
FT	DISULFID	453	471	By similarity.	
FT	DISULFID	648	761	By similarity.	
SQ	SEQUENCE	765 AA;	88369 MW;	E32165421P43E116	CRC64;

Query Match 89.2%; Score 3585; DB 1; Length 765;  
 Best Local Similarity 89.3%; Pred. No. 2.6e-235;  
 Matches 649; Conservative 41; Mismatches 37; Indels 0; Gaps 0;

Qy	13	SRKTYTLTDYLNKTYRLKLSRLWISDHEVLYKQBNILVFNAEYGNSSVPLENSTPDEF	72
Db	38	SRRTYTLADYLNKTYFRMKFYNLRWVSDHEVLYKQBNILFNAYGNSSVPLENSTPDEF	97
Qy	73	GHSINDYSISPDGQFILLENNYVQWRHSYTYASYDIYDLNKRQLITEERIPNNTQWTVWS	132
Db	98	GHSINDYSVSPDRQYILFEYNNVQWRHSYTYASYDIYDLNKRQLITEERIPNNTQWTVWS	157
Qy	133	PVGHKLAYVWNNDIYVKIEBNLPSYRITWTGKEBIIYNGITDWTVEEVFSAYSALWSP	192
Db	158	SVGHKLAYVWNNDIYVKNEPNSQRITWTGKQVIYNGITDWTVEEVFSAYSALWSP	217
Qy	193	NGTFLAYAQFNDTEVPLIEYSFYSDESLOQPKTVRVPKAGAVNPVKFVWNTDSLSS	252
Db	218	NSTFLAYAQFNDTEVPLIEYSFYSDESLOQPKTVKIPYKAGAVNPVKFVWNTDSLSS	277
Qy	253	VTNATSIQITAPASMLIGDHYLCDVTWATERISLQWLRRIQNYSVMDICDYDESSGRWN	312
Db	278	NINATSQIIVPPGSLVIGDHYLCDVTWATERISLQWLRRIQNYSVMDICDYDRSTGRWI	337
Qy	313	CLVARQHIEMSTTCWGRFRPSPHPTLDGNSFKIISNEEGYRHCYFQIDKDCCTPIT	372
Db	338	SSVGRQHIETSTTCWGRFRPSPHPTLDGNSFKIISNEEGYRHCYFQIDKDCCTPIT	397
Qy	373	KGTWEVIGIEALTSIDYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYS	432
Db	398	KGAEVIGIEALTSIDYLYISNEYKMGPGGRNLYKIQLNDYTKVTCLSCELNPERCOYS	457
Qy	433	VSPSKEAKYYQLRCSGPGLPLYTLHSSVNDKGLRVLEDNSALDKQLQNVQMPKCLDPII	492
Db	458	VSPSQEAKYYQLRCSGPGLPLYTLHSSVNDKGLRVLEDNSALDKQLQNVQMPKCLDPIH	517
Qy	493	LNETHKFWYQMLPPHFDKSKYPLLLDVYAGPCSKADTVFRLNWTATLASTENIIVASP	552
Db	518	LHGTKFWYQMLPPHFDKSKYPLLEEVYAGPCSKADAI FRLNWTATLASTENIIVASF	577
Qy	553	DGRSGYQGGKIMHAINRRLTGTFEVEDQIRAAQFQSKMGFVDNKRKRIAIWGSYGYVTSM	612
Db	578	DGRSGYQGGKIMHAINRRLTGTFEVEDQIEATQFQSKMGFVDNKRKRIAIWGSYGYVTSM	637
Qy	613	VLGSGGVFKCGIAGVAPVSRWEYDYSVYTERYMGLEPTPEDNLDHYRNSTVMSRAENKQV	672
Db	638	VLGSGGVFKCGIAGVAPVSKWEYDYSVYTERYMGLEPTPEDNLDYSRNSTVMSRAENKQV	697
Qy	673	EYLLIHGTADDNVHFQOQSAQISKALVDVGVDFQAWTTTDEHDGIIASTAHQHIYTHSHF	732
Db	698	EYLLIHGTADDNVHFQOQSAQISKALVDAGVDFQSMWTTTDEHDGIIASTAHQHIYTHSHF	757
Qy	733	IKOCFSL 739	
Db	758	LKOCFSL 764	

RESULT 5

DPP4\_FELCA  
ID DPP4\_FELCA STANDARD; PRT; 765 AA.  
AC Q9NZI7;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP IV) (T-cell activation antigen CD26) (Contains: dipeptidyl peptidase 4 membrane form (Dipeptidyl peptidase IV membrane form); Dipeptidyl peptidase 4 soluble form (Dipeptidyl peptidase IV soluble form)).  
DE Name=DPP4; Synonyms=CD26;  
GN Felis silvestris catus (Cat).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae; OC Felinae; Felis.  
OX NCBI\_TaxID=9685;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Periphereal blood;  
RX MEDLINE=20094000; PubMed=10630304; DOI=10.1007/s002510050616;  
RA Nishimura Y., Miyazawa Y., Ikeda Y., Izumiya Y., Nakamura K., Sato E., Mikami T., Takahashi E.;  
RT "Molecular cloning and sequencing of a cDNA encoding the feline T-cell activation antigen CD26 homologue.";  
RL Immunogenetics 50:366-368(1999).  
CC -!- FUNCTION: Removes N-terminal dipeptides sequentially from polypeptides having unsubstituted N-termini provided that the penultimate residue is proline (By similarity).  
CC -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided Zaa is neither Pro nor hydroxyproline.  
CC -!- SUBUNIT: Homodimer (By similarity).  
CC -!- SUBCELLULAR LOCATION: type II membrane protein. Also exists in a soluble form (By similarity).  
CC -!- PTM: The soluble form (SDPP) derives from the membrane form (MDPP) by proteolytic processing (By similarity).  
CC -!- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.  
CC -----  
DR EMBL; AB023952; BAA92344.1; -; mRNA.  
DR HSP; P27487; LPFG.  
DR SMR; Q9NZI7; 38-765.  
DR MEROPS; S09.003; -.  
DR InterPro; IPR002471; Pept S9 AS.  
DR InterPro; IPR001375; PeptIdase S9.  
DR InterPro; IPR002469; PeptIdase\_S9B.  
DR InterPro; IPR000379; Ser\_estr.  
DR Pfam; PF00930; DPPIV\_N; 1.  
DR Pfam; PF00326; PeptIdase S9; 1.  
DR PROSITE; PS00708; PRO\_ENDOPEP\_S9; 1.  
KW Aminopeptidase; Glycoprotein; Hydrolase; Protease; Serine protease;  
KW Signal-anchor; Transmembrane.  
FT CHAIN 1 765 Dipeptidyl peptidase 4 membrane form.  
FT CHAIN 38 765 Dipeptidyl peptidase 4 soluble form (By similarity).  
FT TOPO\_DOM 1 6 Cytoplasmic (Potential).  
FT TRANSMEM 7 29 Signal-anchor for type II membrane protein (Potential).  
FT TOPO\_DOM 30 765 Extracellular (Potential).  
FT ACT\_SITE 629 629 Charge relay system (By similarity).  
FT ACT\_SITE 707 707 Charge relay system (By similarity).  
FT ACT\_SITE 739 739 Charge relay system (By similarity).  
FT CARBOHYD 84 84 N-linked (GlcNAc...) (By similarity).  
FT CARBOHYD 91 91 N-linked (GlcNAc...) (By similarity).  
FT CARBOHYD 149 149 N-linked (GlcNAc...) (By similarity).  
FT CARBOHYD 178 178 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 228 228 N-linked (GlcNAc...) (By similarity).  
FT CARBOHYD 280 280 N-linked (GlcNAc...) (By similarity).

FT CARBOHYD 320 320 N-linked (GlcNAc...) (By similarity).  
FT CARBOHYD 330 330 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 331 331 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 519 519 N-linked (GlcNAc...) (By similarity).  
FT CARBOHYD 684 684 N-linked (GlcNAc...) (By similarity).  
FT DISULFID 384 393 By similarity.  
FT DISULFID 443 446 By similarity.  
FT DISULFID 453 471 By similarity.  
FT DISULFID 648 761 By similarity.  
SQ SEQUENCE 765 AA; 88213 MW; 3EFCB98A22B175D9 CRC64;  
Query Match 88.6%; Score 3560; DB 1; Length 765;  
Best Local Similarity 88.0%; Pred. No. 1.3e-233;  
Matches 641; Conservative 47; Mismatches 40; Indels 0; Gaps 0;  
QY 13 SRKTYTLTYLKYVRLKLYSLRWISDHEYLKQENNVNNAEYVNGSVLENSTFDEF 72  
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 97  
QY 73 GHSINDYISDPGQFILLEVNVKQWRHSYASDIYDLNKKQLITEERIPNNTQWTVS 132  
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 157  
QY 133 PVGHKLAYVMNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEVSALWSP 192  
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 217  
QY 139 NGTFILAYAQFNDTEVPLIEYSFYSDLSQYKTVRPYKAGAVNPTVKFFVNTDSLSS 252  
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 277  
QY 218 KGTFLAYAQFNDTQVPLIEYSFYSDLSQYKTVRPYKAGAVNPTVKLVFKTDNLNP 312  
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 337  
QY 253 VTNATSIOTAPASMLIGDHYLCDVTWATOBRIQLWRRIQNSVMDICDYDESSGRWN 372  
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 397  
QY 313 CLVARQHIEMSTGWVGRFRPSEPHFTLDGNSFYKIIISNEGYRHCYFQIDKKDCTFTT 432  
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 457  
QY 373 KGTWEVIGIEALTSYLYYISNEYKMGPGRNLYKIQLSDYTKVTCSCELNPERCOYVS 492  
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 517  
QY 433 VSFSEAKYQYLRCSGGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVMPSSKLDFTI 552  
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 577  
QY 493 LNETKFWYQMLPPHFDKSKYPLLLDVYAGCSQKADTVFRLNWTYLASTENIIVASF 612  
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 637  
QY 553 DGRSGYOGDKIMHAINRLCTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGVVTSM 672  
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 697  
QY 578 DGRSGYOGDKIMHAINRLCTFEVEDQIEAARQFSKMGFVDDKRIATWGSYGGVVTSM 732  
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 757  
QY 613 VLGSYGKVFYKCGIAVAPVSRWEYDSVYTERYMGFLPTPEDNLDHYRNSVTMSRAENFKQV 772  
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 797  
QY 638 VLGSYGKVFYKCGIAVAPVSRWEYDSVYTERYMGFLPTPDNDLDYKNSVTMSRAENFKQV 822  
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 847  
QY 673 EYLLIHGTADDNVHFQQAISKALVDVGVDFQAMWYTTDEDHGIIASSTAHQHIYTHMSHF 922  
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 947  
QY 698 EYLLIHGTADDNVHFQQAISKALVDAGVDFOAMWYTTDEDHGIIASGPAPHQHIYTHMSHF 1022  
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1047  
QY 733 IKQCFSLP 740  
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 765  
QY 758 IKQCFSLP 765  
RESULT 6  
DPP4\_FIG  
ID DPP4\_FIG  
AC P22411; Q866G2;  
DT 01-AUG-1991 (Rel. 19, Created)



```

Db      519 LGHTKEWYQMLPPHFDKSKKYPDLLIEVAGCSQKVDIVFRLSWATYLASTENIIVASF 578
Qy      553 DGRGSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNKRITAIWGSYGGVYTSM 612
Db      579 DGRGSGYQGDKIMHAINRRLGTFEVEDQIEATRQFSKMGFVDDKRIAIWGSYGGVYTSM 638
Qy      613 VLGGSGVPCGKTAVAPVRWEYDSVYTERYMGLETPEDNLDHYRNSVTMSRAENFKQV 672
Db      639 VLGGSGVPCGKTAVAPVRWEYDSVYTERYMGLETPEDNLDHYRNSVTMSRAENFKQV 698
Qy      673 EYLLIHGTADDNVHFQSQAKISKALVDGVDFQAMWYTDDEHGIASTAHQHIYTHMSHF 732
Db      699 EYLLIHGTADDNVHFQSQAKISKALVDAGVDFQTMWYTDDEHGIASTAHQHIYTHMSHF 758
Qy      733 IKQCFSLP 740
Db      759 LKQCFSLP 766

RESULT 7
DPP4 RAT
ID      DPP4 RAT      STANDARD;      PRT;      767 AA.
AC      P14740;
DT      01-APR-1990 (Rel. 14, Created)
DT      01-FEB-2005 (Rel. 46, Last sequence update)
DE      Dipeptidyl peptidase 4 (SC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP
DE      IV) (T-cell activation antigen CD26) (GP110 glycoprotein) (Bile
DE      canaliculus domain-specific membrane glycoprotein) (Contains:
DE      Dipeptidyl peptidase 4 membrane form (Dipeptidyl peptidase IV membrane
DE      form); Dipeptidyl peptidase 4 soluble form (Dipeptidyl peptidase IV
DE      soluble form); Dipeptidyl peptidase 4 60 kDa soluble form (Dipeptidyl
DE      peptidase IV 60 kDa soluble form)).
DE      Name=Dpp4; Synonyms=CD26;
GN      Rattus norvegicus (Rat).
OS      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC      Muroidae; Muridae; Murinae; Rattus.
OX      NCBI_TaxID=10116;
RN      [1]
RP      NUCLEOTIDE SEQUENCE, AND PARTIAL PROTEIN SEQUENCE.
RX      MEDLINE=89123496; PubMed=2563382;
RA      Ogata S., Misumi Y., Ikehara Y.;
RT      "Primary structure of rat liver dipeptidyl peptidase IV deduced from
RT      its cDNA and identification of the NH2-terminal signal sequence as the
RT      membrane-anchoring domain.";
RL      J. Biol. Chem. 264:3596-3601(1989).
RN      [2]
RP      NUCLEOTIDE SEQUENCE.
RX      MEDLINE=88068516; PubMed=3479775;
RA      Hong W., Doyle D.;
RT      "cDNA cloning for a bile canaliculus domain-specific membrane
RT      glycoprotein of rat hepatocytes.";
RL      Proc. Natl. Acad. Sci. U.S.A. 84:7962-7966(1987).
RN      [3]
RP      NUCLEOTIDE SEQUENCE OF 1-40.
RX      MEDLINE=89034185; PubMed=3182821;
RA      Hong W.J., Doyle D.;
RT      "Membrane orientation of rat gp110 as studied by in vitro
RT      translation.";
RL      J. Biol. Chem. 263:16992-16998(1988).
RN      [4]
RP      PROTEIN SEQUENCE OF 28-58, AND TISSUE SPECIFICITY.
RX      MEDLINE=90228896; PubMed=1970322;
RA      McCaughan G.W., Wicks J.E., Creswick P.F., Gorrell M.D.;
RT      "Identification of the bile canaliculus cell surface molecule GP110 as
RT      the ectopeptidase dipeptidyl peptidase IV: an analysis by tissue
RT      distribution, purification and N-terminal amino acid sequence.";
RL      Hepatology 11:534-544(1990).
RN      [5]
RP      PROTEIN SEQUENCE OF 281-302, AND MUTAGENESIS OF GLY-629; TRP-630;
RP      SER-631; TRY-632 AND GLY-633.

```

```

RC      TISSUE=Kidney;
RX      MEDLINE=94128239; PubMed=7905271;
RA      Iwaki-Egawa S., Watanabe Y., Fujimoto Y.;
RT      "N-terminal amino acid sequence of the 60-kDa protein of rat kidney
RT      dipeptidyl peptidase IV.";
RL      Biol. Chem. Hoppe-Seyler 374:973-975(1993).
RN      [6]
RP      PROTEIN SEQUENCE OF 624-648.
RX      MEDLINE=92190188; PubMed=1347701;
RA      Ogata S., Misumi Y., Tsuji E., Takami N., Oda K., Ikehara Y.;
RT      "Identification of the active site residues in dipeptidyl peptidase IV
RT      by affinity labeling and site-directed mutagenesis.";
RL      Biochemistry 31:2582-2587(1992).
RN      [7]
RP      SIGNAL-ANCHOR.
RX      MEDLINE=90338089; PubMed=1974258; DOI=10.1083/jcb.111.2.323;
RA      Hong W., Doyle D.;
RT      "Molecular dissection of the NH2-terminal signal/anchor sequence of
RT      rat dipeptidyl peptidase IV.";
RL      J. Cell Biol. 111:323-328(1990).
CC      -!- FUNCTION: Removes N-terminal dipeptides sequentially from
CC      polypeptides having unsubstituted N-termini provided that the
CC      penultimate residue is proline.
CC      -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC      Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
CC      Zaa is neither Pro nor hydroxyproline.
CC      -!- SUBUNIT: Homodimer.
CC      -!- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a
CC      soluble form.
CC      -!- TISSUE SPECIFICITY: Expressed in bile ducts and other epithelial
CC      brush borders (small intestine, kidney, colon, pancreatic duct);
CC      acinar structures in salivary glands; endothelial structures and T
CC      cell areas in thymus, spleen and lymph node.
CC      -!- PTM: The soluble form (SDPP) derives from the membrane form (MDPP)
CC      by proteolytic processing.
CC      -!- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
CC
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use as long as its content is in no way modified and this statement is not
CC      removed.
CC
CC      EMBL; J04591; AAA41096.1; -; mRNA.
CC      EMBL; J02997; AAA41272.1; -; mRNA.
CC      PIR; A39914; A39914.
CC      HSSP; P27487; 1PFQ.
CC      SMR; P14740; 37-766.
CC      MEROPS; S09.003; -.
CC      Ensembl; ENSRNOG0000030763; Rattus norvegicus.
CC      RGD; 2515; Dpp4.
CC      GO; GO:0005624; C:membrane fraction; IDA.
CC      GO; GO:0005625; C:soluble fraction; IDA.
CC      GO; GO:0004274; F:dipeptidyl-peptidase IV activity; TAS.
CC      GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
CC      InterPro; IPR002471; Pept_S9_AS
CC      InterPro; IPR001375; Peptidase_S9.
CC      InterPro; IPR002469; Peptidase_S9B.
CC      InterPro; IPR000379; Ser_estrs.
CC      Pfam; PF00930; DPPIV_N; I.
CC      Pfam; PF00326; Peptidase_S9; 1.
CC      PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW      Aminopeptidase; Direct protein sequencing; Glycoprotein; Hydrolase;
KW      Protease; Serine protease; Signal-anchor; Transmembrane.
FT      CHAIN      1      767
FT      CHAIN      37      767
FT      CHAIN      281      767
FT
FT      TOPO_DOM      1      6
FT      TRANSMEM      7      28
FT
FT      TOPO_DOM      29      767
FT      ACT_SITE      631      631

```

FT	ACT_SITE	709	709	Charge relay system (By similarity).
FT	ACT_SITE	741	741	Charge relay system (By similarity).
FT	CARBOHYD	83	83	N-linked (GlcNAc. .) (By similarity).
FT	CARBOHYD	90	90	N-linked (GlcNAc. .) (By similarity).
FT	CARBOHYD	148	148	N-linked (GlcNAc. .) (By similarity).
FT	CARBOHYD	217	217	N-linked (GlcNAc. .) (By similarity).
FT	CARBOHYD	227	227	N-linked (GlcNAc. .) (By similarity).
FT	CARBOHYD	319	319	N-linked (GlcNAc. .) (By similarity).
FT	CARBOHYD	521	521	N-linked (GlcNAc. .) (By similarity).
FT	CARBOHYD	686	686	N-linked (GlcNAc. .) (By similarity).
FT	DISULFID	383	395	By similarity.
FT	DISULFID	445	448	By similarity.
FT	DISULFID	455	473	By similarity.
FT	DISULFID	650	763	By similarity.
FT	MUTAGEN	629	629	G->A: Reduced activity.
FT	MUTAGEN	629	629	G->R: Reduced activity.
FT	MUTAGEN	630	630	W->R: No effect on activity.
FT	MUTAGEN	631	631	S->A: Reduced activity.
FT	MUTAGEN	632	632	Y->F: No effect on activity.
FT	MUTAGEN	632	632	Y->G: Reduced activity.
FT	MUTAGEN	632	632	Y->L: Reduced activity.
FT	MUTAGEN	633	633	G->A: Reduced activity.
FT	MUTAGEN	633	633	G->S: Reduced activity.
FT	CONFLICT	38	38	R -> A (in Ref. 1).
FT	CONFLICT	54	54	Missing (in Ref. 4).
FT	CONFLICT	183	183	I -> T (in Ref. 2).
FT	CONFLICT	332	332	T -> N (in Ref. 2).
FT	CONFLICT	352	352	C -> V (in Ref. 2).
FT	CONFLICT	394	394	V -> D (in Ref. 2).
FT	CONFLICT	562	562	L -> F (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT	624	624	R -> Q (in Ref. 2).
FT	CONFLICT			

Query Match  
 Best Local Similarity 84.8%; Score 3408.5; DB 1; Length 767;  
 Matches 618; Conservative 53; Mismatches 56; Indels 3; Gaps 1;

Qy	13	SRKTYTLDYLNKTYRLKLSLRWISDHEYLKQENNILVFNABYSGNSVFLENSTDEF	72
Db	37	SRRTYTLADYLNKTYFRVKSLSRWSDSEYLYKQENNILVFNABYSGNSVFLENSTDEF	96
Qy	73	GHSINDYSISDGGFILLVYVYQWRHSYTSASDIYDLNKRQLITRIPNTQWTWS	132
Db	97	GDSISDYSVSDRLFLVLLVYVYQWRHSYTSASDIYDLNKRQLITRIPNTQWTWS	156
Qy	133	PVGHKLAVWNDIYVKEPLNPSYRIITWTKEDIYNGITDWWYVEVRSAYSALWSP	192
Db	157	QEGHKLAVWNDIYVKEPLNPSYRIITWTKEDIYNGITDWWYVEVRSAYSALWSP	216
Qy	193	NGTFLAYAQFNDTEVPLIEXSYFSDLSIQYPTKTVRVPYPKAGAVNPTVKPFVNVNTSLSS	252
Db	217	NGTFLAYAQFNDTEVPLIEXSYFSDLSIQYPTKTVRVPYPKAGAVNPTVKPFVNVNTSLSS	276
Qy	253	VTNATSIQTAPASMLIGDHYLVDVWATQBRISLOWLRRIQNTVMDICDYDESSGRWN	312
Db	277	TTTTIPMQITAPASVTTGDHYLVDVWATQBRISLOWLRRIQNTVMDICDYDESSGRWN	336
Qy	313	CLVAROHEIEMSTTCWGRFEPSEPHFTLDGNSFYKILISNBERGRIHCVFQIDKDD---CT	369
Db	337	CPTTQEHIEFISATCGCRFRPAEHPFTSDGSGSFYKIVSDKDGKHIQFQKDRKPEQVCT	396
Qy	370	FITKGTWEVIGIEALTSIDLYIISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQ	429
Db	397	FITKGAWEVISIEALTSIDLYIISNEYKMGPGGRNLKYIQLTDHTNKKCLSCDLNPERCQ	456
Qy	430	YYSVSPFEAKYIQLRCGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKLD	489
Db	457	YYSVSLSEAKYIQLRCGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKLD	516
Qy	490	FIILNETKFWQMTLLPHEFDSKCYKPLLDVYAGPCSKADTVPRLNWTATLASTENIIV	549
Db	517	FIVLNETRFWQMTLLPHEFDSKCYKPLLDVYAGPCSKADTVPRLNWTATLASTENIIV	576

Qy	550	ASFDGRSGYQGDKIMHAINRRLCTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGTV	609
Db	577	ASFDGRSGYQGDKIMHAINRRLCTFEVEDQIEAARQFLKMGFVDSKRVAIWGSYGGTV	636
Qy	610	TSWVLGSGSGVFKCGIAVAPVSRWEYVDSVYTERYMGILPTPEDNLHDYRNSTVMSRAENF	669
Db	637	TSWVLGSGSGVFKCGIAVAPVSRWEYVDSVYTERYMGILPTPEDNLHDYRNSTVMSRAENF	696
Qy	670	KQVEYLLHGTADDNVHFQSSAQISKALVDVGVDFQAMWYTDDEHGIIASSTAHQHIYTHM	729
Db	697	KQVEYLLHGTADDNVHFQSSAQISKALVDVGVDFQAMWYTDDEHGIIASSTAHQHIYTHM	756
Qy	730	SHFIKQCFSL 739	
Db	757	SHFIKQCFSL 766	

## RESULT 8

ID	DPP4_MOUSE	STANDARD;	PRT;	760 AA.
AC	P28843;			
DT	01-DEC-1992 (Rel. 24, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	13-SEP-2005 (Rel. 48, Last annotation update)			
DE	Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP IV) (T-cell activation antigen CD26) (Thymocyte-activating molecule) (THAM) [Contains: Dipeptidyl peptidase 4 membrane form (Dipeptidyl peptidase IV membrane form); Dipeptidyl peptidase 4 soluble form (Dipeptidyl peptidase IV soluble form)].			
GN	Name=Dpp4; Synonyms=Cd26;			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;			
OC	Muroidea; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	STRAIN=Swiss; TISSUE=Thymus;			
RC	MEDLINE=92129288; PubMed=1370813;			
RA	Marguet D.A., Bernard A.-M., Vivier I., Darmoul D., Naquet P.,			
RA	Pierres M.;			
RT	"cDNA cloning for mouse thymocyte-activating molecule. A			
RT	multifunctional ecto-dipeptidyl peptidase IV (CD26) included in a			
RT	subgroup of serine proteases.";			
RT	J. Biol. Chem. 267:2200-2208(1992).			
RN	[2]			
RP	SEQUENCE REVISION.			
RL	Marguet D.A.;			
RL	Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	NUCLEOTIDE SEQUENCE.			
RC	STRAIN=B10.A; TISSUE=Liver;			
RC	MEDLINE=95092780; PubMed=7999781;			
RA	Bernard A.-M., Mattei M.-G., Pierres M., Marguet D.;			
RT	"Structure of the mouse dipeptidyl peptidase IV (CD26) gene.";			
RT	Biochemistry 33:15204-15214(1994).			
RN	[4]			
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].			
RC	STRAIN=C57BL/6J; TISSUE=Kidney;			
RC	MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;			
RA	Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,			
RA	Nikaido I., Osato N., Saigo R., Suzuki H., Yamanaka I., Kiyosawa H.,			
RA	Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojohori T.,			
RA	Baldarelli R., Hilli D.P., Bult C., Hume D.A., Quackenbush J.,			
RA	Schirni L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,			
RA	Blake J.A., Brad D., Brusic V., Chothia C., Corbani L.E., Cousins S.,			
RA	Dalla E., Dragani T.A., Fletcher C.P., Forrest A., Frazer K.S.,			
RA	Gaasterland T., Gariboldi M., Giesi C., Godzik A., Gough J.,			
RA	Grimmond S., Guatinchich S., Hirokawa N., Jackson I.J., Jarvis E.D.,			
RA	Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,			
RA	Konagaya A., Kurochkin I.V., Lee Y., Lennhard B., Lyons P.A.,			
RA	Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,			
RA	Nagashima T., Numata K., Okido T., Pavan W.J., Portea G., Resole G.,			





Db 153 PEGHKLAVWKNDIIVKVEPHLPSPHRTSTGSENVYNGITDWNVEEVEFGAYSALWNSP 212  
QY 193 NGTFLAYAQFNDTEVPLIEYSFYSDLSQYPTVRVPYKAGAVNPTVKFFVYVNTDSLSS 252  
Db 213 NNTFLAYAQFNDTEVPLIEYSFYSDLSQYPTVRVPYKAGAVNPTVKFFVYVNTDSLSS 272  
QY 253 VTNATSIQITAPASMLGDHYLCVDTWATOEIRISLOWLRRIQNYSMVDCDYDESSGRWN 312  
Db 273 SSSAAPIQIPASVARGDHYLCVDTWATEERISLOWLRRIQNYSMVDCDYDKINLTWN 332  
QY 313 CLVARQHEMSTTGWGRFRPSEPHTLDGNSFYKIIISNEEGYRHICVFOIDKDOCTIT 372  
Db 333 CPSEQOHEMSTTGWGRFRPSEPHTSDGSSFYKIIISDQYKXICHFPDKDOCTIT 392  
QY 373 KGTWEVIGIEALTSYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS 432  
Db 393 KGAMEVIGIEALTSYLYISNOYKEMPGGRNLYKIQLTDHTNVKCLSCDLNPERCQYYA 452  
QY 433 VSPSKEAKYQLCRSGPGLPYTLHSSVNDKGLRVLEDSALDMLQNVQMPSKKLDPII 492  
Db 453 VSPSKEAKYQLCGCGPGLPYTLHRSQDHLKELRVLEDSALDMLQNVQMPSKKLDPIV 512  
QY 493 LNETKFWQMLPPHFDKSKYPLLLDVYAGPCOKADTVFRLNWTYLASTENIIVASF 552  
Db 513 LNETKFWQMLPPHFDKSKYPLLLDVYAGPCOKADTVFRLNWTYLASTENIIVASF 572  
QY 553 DGRSGYQGDKIMHAINRRLGTFFVEDQIEAARQFSKMGFVNDKRIAIWGSYGYVTSM 612  
Db 573 DGRSGYQGDKIMHAINRRLGTFFVEDQIEAARQFSKMGFVDSKEVALWGSYGYVTSM 632  
QY 613 VLGSYGVPKCGIAVAPVSRWEYDVSVTERYMGILPTPEDNLDHYNSTVMSRAENPKQV 672  
Db 633 VLGSYGVPKCGIAVAPVSRWEYDVSVTERYMGILPTPEDNLDHYNSTVMSRAEHFKQV 692  
QY 673 EYLLHGTADDNVHFQQAQSKALVDVGVDFQAWMYTDEHGTASSTAHQHIYTHMSHF 732  
Db 693 EYLLHGTADDNVHFQQAQSKALVDVGVDFQAWMYTDEHGTASSTAHQHIYTHMSHF 752  
QY 733 IKQCFSL 739  
Db 753 LQCFSL 759

RESULT 9  
Q75S83 AGKHB PRELIMINARY; PRT; 751 AA.  
AC Q75S83;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DE Dipeptidylpeptidase 4a (EC 3.4.14.5).  
GN Name=DPP4a;  
OS Agkistrodon halye brevicaudus (Korean siamasa snake) (Gloydus halye brevicaudus).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Lepidotesauria; Squamata; Scleroglossa; Serpentes; Colubroides;  
OC Viperidae; Crotalinae; Gloydus.  
OC NCBI\_TaxID=259325;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=venom gland;  
RA Ogawa Y., Yanooshita R., Murayama N., Higuchi S., Samejima Y.;  
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB158224; BAD06332.1; -; mRNA.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0004177; F:aminopeptidase activity; IEA.  
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.  
DR GO; GO:0016787; F:hydrolase activity; IEA.  
DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR002471; Pept\_S9\_AS.  
DR InterPro; IPR001375; Peptidase\_S9.  
DR InterPro; IPR002469; Peptidase\_S9B.

DR InterPro; IPR000379; Ser\_eastrs.  
DR Pfam; PF00930; DPEIV N; I.  
DR Pfam; PF00326; Peptidase\_S9; 1.  
DR PROSITE; PS00708; PRO\_ENDOPEP\_SER; 1.  
KW Hydrolase; Transmembrane.  
SQ SEQUENCE 751 AA; 86127 MW; 6EA3F2D89BBE97D4 CRC64;  
  
Query Match 62.7%; Score 2520.5; DB 2; Length 751;  
Best Local Similarity 62.4%; Pred. No. 7,9e-163;  
Matches 454; Conservative 114; Mismatches 148; Indels 11; Gaps 5;  
  
QY 14 RKTTLTDLKNTVRLKLYSLRWISDHLYLKQENNLVFNABYGNSSVLENSFTDPERG 73  
Db 36 RKFSLELDYSDPEQKSYNLRWMSGHEVYVYTNQNVLLYINIDRRESIVLSDNLTDSFN 95  
QY 74 HSIINDYSPDQPTLLLEYNVVKWRHSYTSYDIYDLNKRQLITEERIPNNTQVWTSVP 133  
Db 96 SS-QALLSPDKFALLQYSYKVRHSYTSYHYIYDLNNTKTIENPLPTNIQIYSWSP 153  
QY 134 VGHKLAVYVNNNDIIVKIEPNLPSYRITWTGKEDIYNGITDWNVEEVEFSYALSALWNSP 193  
Db 154 VGHKLAVYVYRNVVVKATPNASPVQITENGAENKILGLADWVYEEMFGTHSALWNSP 213  
QY 194 GTFLAYAQFNDTEVPLIEYSFYSDLSQYPTVRVPYKAGAVNPTVKFFVYVNTDSLSSV 253  
Db 214 GRFLAFABEINDTEVPVMEYSFYSDLTQYPTKIKIPYKAGAINPTIRLFLVDI----SL 269  
QY 254 TNATSIQITAPASMLGDHYLCVDTWATOEIRISLOWLRRIQNYSMVDCDYDESSGRNC 313  
Db 270 SPKNISEIVAPSSIIISGDHYLSAVTWVTDVTCVQWLRRIQNFVLTICDY---SGAWHC 326  
QY 314 LVARQHIEMSTTGWGRFRPSEPHTLDGNSFYKIIISNEEGYRHICVFOIDKDOCTFITK 373  
Db 327 PKREHLEESKTNVGRFQFSEPYFTSDKISYRIISDSEGYKILHYTDSAGK-VKPTIS 385  
QY 374 GTWEVIGIEALTSYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSV 433  
Db 386 GKWEVISISAVTNNSLYFISNEPEGRPGGRHLYKVDLKNLKKIKITCNSKEAACQYFSV 445  
QY 434 SPSKEAKYQLCRSGPGLPYTLHSSVNDKGLRVLEDSALDMLQNVQMPSKKLDPIIL 493  
Db 446 SFTSDSYKLNLCYGPDLPTFLQNSITDKAIKLTEDNNNNKVLKEIQMPCRLSNITL 505  
QY 494 NETKFWQMLPPHFDKSKYPLLLDVYAGPCOKADTVFRLNWTYLASTENIIVASF 553  
Db 506 HGQTYWQMLPPNFDSEKYPPLLDVYAGPCOKADAAFRINSTYLASSEGIIVASF 565  
QY 554 GRSGYQGDKIMHAINRRLGTFFVEDQIEAARQFSKMGFVNDKRIAIWGSYGYVTSMV 613  
Db 566 GRSGYQGDKILHAIYRRLGTFFVEDQISAAKLFSEMSFVDKRIAIWGSYGYVTSMV 625  
QY 614 LGSGGVKPGCGIAVAPVSRWEYDVSVTERYMGILPTPEDNLDHYNSTVMSRAENPKOVE 673  
Db 626 LGASDVFKCGIAVAPVSRWQYDYSIYTERYMGILPEKKNLNLNFENSTVMSRAENKFRVTD 685  
QY 674 YLLHGTADDNVHFQQAQSKALVDVGVDFQAWMYTDEHGTASSTAHQHIYTHMSHF 733  
Db 686 YLLHGTADDNVHFQQAQSKALVDVGVDFQAWMYTDEHGTASSTAHQHIYTHMSHF 744  
QY 734 KQCFSLP 740  
Db 745 KQCFSLP 751  
  
RESULT 10  
Q75S82 AGKHB PRELIMINARY; PRT; 751 AA.  
ID Q75S82 AGKHB PRELIMINARY; PRT; 751 AA.  
AC Q75S82;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Dipeptidylpeptidase 4b (EC 3.4.14.5).  
GN Name=DPP4b;

[illegible]

Db	626	LGAGSDVFKCGIAVAPRSRQWYDSIYYTERWGLPEKNDNLNFYENSTWAKGNFRITVD	
Qy	674	YLLIHGTADDNVHFPQQAQISKALVDGVDFQAMMYTDDHGIIASSTAQHIIYTHMSHPF	
Db	686	YLLIHGTADDNVHFPQQAQISKALVAEDVDFQAMMYTKDHGIGGH-AHSHIYQHMSPM	
Qy	734	KQCFSLP 740	
Db	745	KQCFKL 751	
 RESULT 11 ID Q5Z181 CHICK PRELIMINARY; PRT; 759 AA.			
AC	Q5Z181;		
DT	25-OCT-2004 (TrEMBLrel. 28, Created)		
DT	25-OCT-2004 (TrEMBLrel. 28, Last sequence update)		
DT	25-OCT-2004 (TrEMBLrel. 28, Last annotation update)		
DE	Hypothetical protein.		
GS	ORFNames=RCJMB04_29g21;		
ON	Gallus Gallus (Chicken).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;		
OC	Gallus.		
OX	NCBI_TaxID=9031;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	SPAIN=CB; TISSUE=Bursa;		
RA	Caldwell R.B., Kierzek A.M., Arakawa H., Bezzubov Y., Zaim J.,		
RA	Fiedler P., Kutner S., Blagodatski A., Kostovska D., Kotter M.,		
RA	Piacchi J., Carinci P., Hayashizaki Y., Buerstede J.M.;		
RT	"Full-length cDNAs from chicken bursal lymphocytes to facilitate		
RT	gene function analysis.";		
RL	Genome Biol. 6:B6-B6 (2005).		
DR	EMBL; AJ270903; CAG32582.1; -; mRNA.		
DR	GO; GO:0016030; C-membrane; IEA.		
DR	GO; GO:0004177; F:aminopeptidase activity; IEA.		
DR	GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.		
DR	GO; GO:0016787; F:hydrolase activity; IEA.		
DR	GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.		
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.		
DR	InterPro; IPR002471; Pept_S9_AS.		
DR	InterPro; IPR001375; Peptidase_S9.		
DR	InterPro; IPR002469; Peptidase_S9B.		
DR	InterPro; IPR000379; Ser_sentr.		
DR	Pfam; PF00930; DPPIV_N; I.		
DR	Pfam; PF00326; Peptidase_S9; 1.		
DR	PROSITE; PS00708; PRO_ENDOPEP_SER; 1.		
KW	Hydrolase; Hypothetical protein; Transmembrane.		
SQ	SEQUENCE 759 AA; 86343 MW; B9A0C38CCB8644AC CRC64;		
 Query Match 62.7%; Score 2519; DB 2; Length 759; Best Local Similarity 64.7%; Pred. No. le-162; Matches 474; Conservative 93; Mismatches 150; Indels 16; Gaps			
Qy	13	SRKYTTLTDLKNTYRLKLYSLRMISDHEYLKQBN-NILVFNAEYGNSSVFLENSTFDE	71
Db	38	SRSYTLLENLYNDVVYKTNHLQWISNGYLETSGNHILRFDAETGTSSVLLNTTISI	97
Qy	72	FGHSINDYSISPDGQIFALLEYNVKQRHSYTASYDIYLNKRQLITEERI PNNTQMTW	131
Db	98	--HEATTAILSPDQRFAQQVKEKLRHSYTASYHYVDFTNTSILLDALLPNDTQYIS	155
Qy	132	SPUGHKLAYWNNDIYKIEPNLSRYLTRWTCKEDIIVNGITDWYVEEVFSAYSALWS	191
Db	156	SPVGHKLAYVMNNIYIKASTAAPVQITSNGEENKIENGIPDWYVEEFMGSHSALWS	215
Qy	192	PNGHTFLAYAQENDTEVPPIEYSFYSDLESIQPKTVRVPYPKAGAVNPVTVKFFVN	251
Db	216	PNGNFVAYAANDTEVPVIESFYSFSEDLTQPKTIIRIYPKAGAKNPVTVKFFIVD--	272
Qy	252	SVTWATSIQTAPASMLIGDHYLCDVTTWATQERISLQWLRRNTQNTVSMDICDYDESSGRW	311

Db 273 MLPDFNSTEISPPABIKSGDHYLSVVTWTERICQLWLRLRIQNTSVLTICDFESATGNW 332  
 Qy 312 NCLVARQH---IEMSTGWGVRFRPSEPHFTLDGNSFYKISINBERGYRHICVFQDKKDC 368  
 Db 333 TC--HRKNFNKVKVLAGW--QISAICLPAPDNTTYKVSNTGYKHIIHYNGTEAPV 388  
 Qy 369 TFIKGTWEVIGIEALTSYLYISNEYKMGPGGRNLYKIQLSDYTKVT-CLSCELNPER 427  
 Db 389 P-ITGKWEVISIAAVTKYFLYISNQNGEMPGGRNLYKMLLESPPKSTQCVCSDLNQR 447  
 Qy 428 COYTSVSSEKAKYQOLRCSPGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVOMPSKK 487  
 Db 448 COYTSASFSKQAYQOLNCLGPGPLMSTLHRSDDQVRLRYLNNTELENSLKDIOIPSKK 507  
 Qy 488 LDFILNTEKWKYQMLPPHFDKSKYPLLLDVYAGPCSQKADTVRLNWNATYLASTENI 547  
 Db 508 LGSITVGGYNLYQMLPPHLDSSKYPLELLEVYAGPCSQKVDHVRFRINWATYLASTEQI 567  
 Qy 548 IVASPDGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDMKRIAINGWSYGG 607  
 Db 568 IVASPDGRSGYQGDIEIMHAINRRLGTVEVEDQISAAFTSEMSFVDKRIAINGWSYGG 627  
 Qy 608 YVTSNVLGSGGVKCGTAVAPVSRWYDYDYTYRYNGLPTPEPNLHYRNSVTMSRAE 667  
 Db 628 YVTSNVLGSGGVKCGTAVAPVSRWQYDYDYTYRYNGLPTPEPNLHYRNSVTMSRAE 687  
 Qy 668 NFKOVEYLLIHTADDNVHFOQSAQISKALVDGVDFQAMWYTDSDHGFIASSTAHOHYT 727  
 Db 688 KFEKEVEYLLIHTADDNVHFOQAQISKALVDAEVDFOAMWYTDKHGI-SQAHKHYYT 746  
 Qy 728 HNSHFIKQCFSLP 740  
 Db 747 HNSHFIKQCFSLP 759

## RESULT 12

Q641D6\_XENLA PRELIMINARY; PRT; 737 AA.  
 AC Q641D6;  
 DT 25-OCT-2004 (TRENBLrel. 28, Created)  
 DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)  
 DE LOC397888 protein.  
 GN Name=LOC397888;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
 OC Xenopodinae; Xenopus; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative.";  
 RL Dev. Dyn. 225:384-391(2002).  
 RP [2]  
 RN NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J.G., Haieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,  
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bobak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield J.S.N., Krzyzinski M.I., Jones S.J.M., Marra M.A.,  
 RA "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Kidney;  
 RA Klein S., Gerhard D.S.;  
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBAJ databases.  
 DR EMBL; BC082401; AA082401.1; -; mRNA.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0004177; F:aminopeptidase activity; IEA.  
 DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.  
 DR GO; GO:0016787; F:hydrolase activity; IEA.  
 DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR002471; Pept\_S9\_AS.  
 DR InterPro; IPR001375; Peptidase S9.  
 DR InterPro; IPR002469; Peptidase S9B.  
 DR InterPro; IPR000379; Ser\_estr.  
 DR Pfam; PF00930; DPIP\_N; I.  
 DR Pfam; PF00326; Peptidase S9; 1.  
 DR PROSITE; PS00708; PRO\_ENDOPEP\_SER; 1.  
 KW Hydrolase; Transmembrane.  
 SQ SEQUENCE 737 AA; 84184 MW; 1E08CF94EDBCC4B CRC64;  
 Query Match 58.4%; Score 2348; DB 2; Length 737;  
 Best Local Similarity 59.0%; Pred. No. 4.3e-151;  
 Matches 430; Conservative 111; Mismatches 168; Indels 20; Gaps 9;  
 Qy 14 RKTVTLTDYLNRYRLKLYSLRWISDHRYLYKQBNILVFNARYGNSSVPLENSTFDFBG 73  
 Db 25 RKTFTLEDFSDYRPRKSPGLKWSSENFVPRDKDNLVIFNVNNTTTEISNTTI---- 80  
 Qy 74 HSIND--YSISPDQFILLENNYKQWRHSYASDIYDLNKRQLITEIRIPNNTQWTV 131  
 Db 81 HNSNSFYTLSEDRKALQYNEKLRHSYASDIYDIKKEIVAANLPLNKIYITW 140  
 Qy 132 SPVGHKLAYVWNNDIYVKIRNPLSPYRIWTGKEDIYNGITDQVYREVFSAYSALWWS 191  
 Db 141 SPVGHKLAYVWNNIYIKVPGGISTITTINGENKILNGIPDWVIEEMFSTVALWWS 200  
 Qy 192 PNGTFLAYAFNDTEVPLIYSFYSDESLOQPKTVRVPYPKAGAVNPTVKPFVNVNDSLS 251  
 Db 201 PDATSLAYVEFNDTDVPIEVSFYGEDSDQYPTHTVPIPKAGARNPTVRLFAVNTRSLA 260  
 Qy 252 SVTNATSIQTAPASMLIGHYLCVDTWATQERISLQWLRRIQNYVNDICDYDESSGRW 311  
 Db 261 VI---NPVEILPPEELRSIDHYISGINWVTDNKMAVQWLRRIQNVSLITMCE---GA 313  
 Qy 312 NCLVARQHIEWSTTGWGVRPSPHPTLDGNSFYKISINBERGYRHICVFQDKKDCFTFI 371  
 Db 314 NC---QPPVYSQSTGWWGVYFQPSAPYFDGLKTYKILISNBERGYRHILFPG- 370  
 Qy 372 TKGTEVIGIEALTSYLYISNEYKMGPGGRNLYKIQLS-SDYTKVTCLSCELPERCQY 430  
 Db 371 TSGNWEVTSIATVASNLFYVSNE--GFPGRRLYKIRLDSGYSYSAQCVCNTCQRCQ 428  
 Qy 431 YSVSFSKAKYQILRCSPGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVOMPSKKLDF 490  
 Db 429 YSAFYSKNSKYSLNCGNPGPLIYTVYVNSSNDRTRTMEDNEDLKLLEDIQMPTKENKS 488  
 Qy 491 IILNETKFWQMLPPHFDKSKYPLLLDVYAGPCSQKADTVRLNWNATYLASTENIIVA 550  
 Db 489 IILDGFLMQLTLPHPHFDKSKYPLLLDVYAGPCSQKADTVRLNWNATYLASTEKLIVA 548  
 Qy 551 SFDGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDMKRIAINGWSYGGYVT 610

```
Db 549 SLDRGSGYQGGKIMHQQIYHKLGTLEVDQDIITAAKHFFSLGFVDPKGMAIWGWSYGGYV 608
QY 611 SMVLGSGGVFKGIAVAPVSRWEYYSYTERYMGLETPEDNLDHYRNSTVMSRAENPK 670
Db 609 SMVLGSGSLFKGIAVAPVSRWYYSYTERYMGLETPEDNLDYLSSTVWARAQKFK 668
QY 671 QVEYLLIHGTADDNVHFOQAISKALVDVGVDFQAMWYTDDEHGTASSTAQHIIYTHMS 730
Db 669 DVEYLLIHGTADDNVHFOQAHSKALVDAQVDFTMWTYTDKDHGI--GGTANRHHYTHMS 727
QY 731 HFIKQCFSL 739
Db 728 HFLKQCFNI 736

RESULT 13
P70092_XENLA PRELIMINARY; PRT; 748 AA.
AC P70092;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Diptericidyl-peptidase IV (EC 3.4.14.5).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Skin;
RX MEDLINE=97390076; PubMed=9249015;
RA Vlaek R., Vilas U., Strobl B., Kreil G.;
RT "cDNA cloning and expression of secreted Xenopus laevis dipeptidyl
aminopeptidase IV.";
RL Eur. J. Biochem. 247:107-113(1997).
DR EMBL; Y08932; CAA70136.1; -, mRNA.
DR HSSP; P27487; IPFQ.
DR MEROPS; S09.003; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004177; F:aminopeptidase activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002471; Pept_S9_AS.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estra.
DR Pfam; PF00930; DPfiv_N; I.
DR Pfam; PF00326; Peptidase_S9; 1.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW Hydrolase; Transmembrane.
SQ SEQUENCE 748 AA; 85587 MW; 31CCF61B25104E5B CRC64;

Query Match 58.0%; Score 2331.5; DB 2; Length 748;
Best Local Similarity 58.8%; Pred. No. 5.8e-150;
Matches 430; Conservative 112; Mismatches 164; Indels 25; Gaps 11;

QY 14 RKTFTLTVDKNTYRLKYLRLWISDHEVLYKQENNLVFNAYEGNVSFLENSTDFEG 73
Db 37 RKTFTLTDFSDYRPKSGLKWSNEFVRDKONVLFNVDNETTTIISNTTI---- 92

QY 74 HSGIND--YGISPDGQFLLILEYVVKQHRHSYASYDIYDLNKRQLITEIRPNNTQWTV 131
Db 93 HNSNSSFYTLSEDRKY-ALQYNEKLRHSYASYHYIDIEKKEIVAANPELNPKIQYITW 151

QY 132 SPVGHKLAVVWNNDIIVKLEPNLPYSRITWTGKEDIYNGITDVTVEEVEFSAJALWS 191
Db 152 SPVGHKLAVVWNNIIVKEVPGGISTITTTNGEHNKILNGIPDWVYEEMFSTNYALWS 211

QY 192 PNGTFLAYAQFNDTEVPLIEYSFYSDESLOQYPKTVRPYPKAGAVNPTVKFFVWNTDSL 251
```

```
Db 212 PDATSLAYFENDTDVPIEYSFYGEDSDQYPHTVTVIPYPRAGARNPTVRLFAVNTSLA 271
QY 252 SVTNATSIQITAPASMLIGDHYLCDVWTATQBRISLQWLRRIQNTYSVMDICDYDESSGRW 311
Db 272 VI---NPVEILLPPBELRSIDHYISGINWVTONTKMAVQMLRRITQNTVSLTMCB---GAAM 324
QY 312 NCL--VARQHIEMSTTGVGRFRPSEPFTLDGNSFYKIIISNEEGVRHICYFQIDKKOCT 369
Db 335 NCQPPVYKQ---STGWGIFQPPAPFTQDLKYKIIISNKEGKXHLHLEF-GSKDPV 379
QY 370 FITKGTWEVIGIEALTSYLYISNEYKMGPGGRNLYKIQL--SDYTKVTCLSCELNPERC 428
Db 380 AITSGNWEVTSIATVASNFLYVSNE--GPGRRQLYKIRLDSGYSACVTCNTRQERC 437
QY 429 QYYSVSFSKEAYIOLRCSPGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKL 488
Db 438 QQYSAYFSKSKYSLNCGPGLPIYTVNSSNDNETRMTMEDNEDLKLLEDIQMPTKEN 497
QY 489 DFIILNETKFWYQMLLPHEDKSKYPLLLDVYAGPCQKADTVFRLNWTATLASTENII 548
Db 498 KSIILIDGFEMLYQLTLPPHPDKSKYPLLLIDVYGPGSKQVDFRLNWTATLASTEKII 557
QY 549 VASFDGRSGYQGGKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKRIAIWGSYGY 608
Db 558 VASLDGRSGYQGGKIMHAIYHKLGTLEVDQDIITAAKHFFSLGFVDPKGMAIWGWSYGY 617
QY 609 VTSVNLGSGGVFKGIAVAPVSRWEYYSYTERYMGLETPEDNLDHYRNSTVMSRAEN 668
Db 618 VTSVNLGSGSLFKGIAVAPVSRWYYSYTERYMGLETPEDNLDYLSSTVWARAQK 677
QY 669 FKQVEYLLIHGTADDNVHFOQAISKALVDVGVDFQAMWYTDDEHGTASSTAQHIIYTH 728
Db 678 FKQVEYLLIHGTADDNVHFOQAHSKALVDAQVDFTMWTYTDKDHGI--GGTANRHHYTH 736
QY 729 MSHFIKQCFSL 739
Db 737 MSHFLKQCFNI 747

RESULT 14
Q53TP5_HUMAN PRELIMINARY; PRT; 760 AA.
AC Q53TP5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein FAP.
GN Name=FAP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Cotton M., Maupin R., Hawkins M., Harkins R.;
RT "The sequence of Homo sapiens BAC clone RP11-576116.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Waterston R.H.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Waterston R.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Wilson R.K.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007750; AAY24205.1; -, Genomic_DNA.
GO; GO:0016020; C:membrane; IEA.
```



RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [6]  
 RP PROTEIN SEQUENCE OF 192-208; 220-240 AND 510-521.  
 RX MEDLINE=94327249; PubMed=7519584;  
 RA Rettig W.J., Su S.L., Fortunato S.R., Scanlan M.J., Raj B.K.M.,  
 RA Garin-Chesa P., Healey J.H., Old L.J.  
 RT "Fibroblast activation protein: purification, epitope mapping and  
 RT induction by growth factors.";  
 RL Int. J. Cancer 58:385-392(1994).  
 CC -!- FUNCTION: May have a role in tissue remodeling during development  
 CC and wound healing, and may contribute to invasiveness in malignant  
 CC cancers.  
 CC -!- CATALYTIC ACTIVITY: Degrades gelatin and heat-denatured type I and  
 CC type IV collagen, but not native type I or type IV collagen. Does  
 CC not cleave laminin, fibronectin, fibrin or casein.  
 CC -!- SUBUNIT: Homodimer, or heterodimer with DPP4. The monomer is  
 CC inactive.  
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Found in cell  
 CC surface lamellipodia, invadopodia and on shed vesicles.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1; Synonyms=L;  
 CC IsoId=Q12884-1; Sequence=Displayed;  
 CC Note=Major isoform;  
 CC Name=2; Synonyms=S, Truncated;  
 CC IsoId=Q12884-2; Sequence=VSP\_005367;  
 CC -!- TISSUE SPECIFICITY: Fibroblast specific.  
 CC -!- INDUCTION: In fibroblasts at times and sites of tissue remodeling  
 CC during development, tissue repair, and carcinogenesis.  
 CC -!- PTM: N-glycosylated.  
 CC -!- PTM: The N-terminus may be blocked.  
 CC -!- SIMILARITY: Belongs to the peptidase S9B family.  
 CC -----  
 CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 CC EMBL; U09278; AAB49652.1; -; mRNA.  
 CC EMBL; U76833; AAC51668.1; -; mRNA.  
 CC EMBL; AF007822; AAF21600.1; -; mRNA.  
 CC EMBL; BC026250; AAH26250.1; -; mRNA.  
 CC PDB; 1268; X-ray; A=39-757, B=-.  
 CC MEROPS; S09.007; -.  
 CC Ensembl; ENSG00000078098; Homo sapiens.  
 CC HGNC; HGNC:3590; PAP.  
 CC MIM; 600403; -.  
 CC GO; GO:0003687; C:integral to plasma membrane; TAS.  
 CC GO; GO:0030027; C:lamellipodium; IDA.  
 CC GO; GO:0005886; C:plasma membrane; NAS.  
 CC GO; GO:0004274; F:dipeptidyl-peptidase IV activity; NAS.  
 CC GO; GO:004222; F:metalloendopeptidase activity; TAS.  
 CC GO; GO:0042803; F:protein homodimerization activity; NAS.  
 CC InterPro; IPR002471; Pept\_S9\_AS.  
 CC InterPro; IPR001375; Peptidase\_S9.  
 CC InterPro; IPR002469; Peptidase\_S9B.  
 CC InterPro; IPR000379; Ser\_estr.  
 CC Pfam; PF00930; DPPIV\_N; 1.  
 CC Pfam; PF00326; Peptidase\_S9; 1.  
 CC PROSITE; PS00708; PRO-ENDOPAPP\_S9; 1.  
 CC 3D-structure; Alternative splicing; Direct protein sequencing;  
 KW Glycoprotein; Hydrolase; Polymorphism; Protease; Serine protease;  
 KW Signal-anchor; Transmembrane.  
 FT TOPO\_DOM 1 4 Cytoplasmic (Potential).  
 FT TRANSMEM 5 25 Signal-anchor for type II membrane  
 FT protein (Potential).  
 FT TOPO\_DOM 26 760 Extracellular (Potential).  
 FT ACT\_SITE 624 624 Charge relay system (By similarity).  
 FT ACT\_SITE 702 702 Charge relay system (By similarity).  
 FT ACT\_SITE 734 734 Charge relay system (By similarity).

FT CARBOHYD	49	49	N-linked (GlcNAc. . .)	(Potential).
FT CARBOHYD	92	92	N-linked (GlcNAc. . .)	(Potential).
FT CARBOHYD	99	99	N-linked (GlcNAc. . .)	(Potential).
FT CARBOHYD	314	314	N-linked (GlcNAc. . .)	(Potential).
FT CARBOHYD	679	679	N-linked (GlcNAc. . .)	(Potential).
FT VARSPLIC	1	521	Missing (in isoform 2). /FTid=VSP_005367.	
FT VARIANT	354	354	R -> T (in dSNP:1126507). /FTid=VAR_019691.	
FT CONFLICT	207	207	P -> A (in Ref. 2 and 5).	
FT CONFLICT	229	229	K -> T (in Ref. 2 and 5).	
SQ SEQUENCE	760 AA;	87821 MW;	AD034B4801BE07EA CRC64;	

Query Match 53.9%; Score 2168; DB 1; Length 760;  
 Best Local Similarity 52.5%; Pred. No. 8.1e-13;  
 Matches 389; Conservative 138; Mismatches 200; Indels 14; Gaps 7;

QY	3	PGSHHHHHHGRKTYTLDYLNKTVRLKLYSLRWISDHELYLKYQ-ENNLVFNABYGNSS	61
DB	28	PSRVHSENTMRALTLDKILNGTFSTYKTFPPNWSIQEYLHQSADNNIVLYNIETGQSY	87
QY	62	VLENSTFDEFGHSIN--DYSISPDGQFILLIENYVVKWRHSYTSYDIYDLNKEQLITE	119
DB	88	TILSNRTM---KSVNASNYGLSPDRQFVLYESDKLWRSYATYTYIYDLSNGEFVRG	143
QY	120	ERIPNNTQWTVSPVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEE	179
DB	144	NELPRPIQYLCWSPVGSKLAVYQNNIYLKQRPDGPQITFNGRENKIFNGIPDWYEE	203
QY	180	EVFSAYSALWSPNGTFLAYAQFNDTEVPLIEYSYSDIESLOYPKTVRVPPKAGAVNPT	239
DB	204	EMLPTKYALWSPNGKFLAYAEFNKDIPVIAYSYYGDE--QYPTINIPYKAGAKNPV	261
QY	240	VKFFVWNTDSLSSVNTATSIQITAPASMLIGHYLCVDTWTATQERISLQWLRRINYSVM	299
DB	262	VRIFITDTPYAYGPGQ---EVPVPMATSSDYFSLWLTWTVDERVCQWLKRVQNVSL	318
QY	300	DICDYDESSGWNCLVARQHIEMSTTGWGRFRPSEPHEFTLDGNSFYKIIISNEEYRHC	359
DB	319	SICDPREDQWQDCEPKTQEHIEESTGTWAGGFFVSRPVFSYDAISYKIFSDKGYKHLH	378
QY	360	YFQIDKDCDCTITKGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDY-TKVTC	418
DB	379	YIKDTVENAIQITSGKEAINIFRVTDLSFLSYSSNEFEYEPGRNIRIISIGSYPPSKK	438
QY	419	LSCELNPERCOYYSVSFSKEAKYIOLRCGCLPLYTLHSSVNDKGLRVLEDNSALDKWL	478
DB	439	VTCHLRKERCQYTTASFADYAKYALCYGPGPISTLHDGTDQEIKILEENKLENAL	498
QY	479	QNVQMPSKKLDPIILNETKFWQMLPPHFDKSKKYPLLLDYYAGPCSKADTVFRLNWA	538
DB	499	KNIQLPKEIKKLEVDLTLWYKMLPPQFDRSKYPLLIQYGGPCSQSVSRVFAVNI	558
QY	539	TYLASTENIIVASFGRSGSGQDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKKI	598
DB	559	SYLASKEGWTALVDGRGTAFQDQKLYAVYKLVGVEVEDQITAVRKFIENGFIDEKRI	618
QY	599	AIWGSYGGYVTSWVLGSGSGVFKGIAVAPVSRMEYVDSVYTERYMGILPTEDNLHVR	658
DB	619	AIWGSYGGYVSSSLALASGTGLFKGIAVAPVSSMEYASVYTERPMGLPTKDDNLHVK	678
QY	659	NSTVMSRAENFKQVEYLIIHGTTADDNVHVFQSAQISKALVDVGVDFQAMWYTDDEHGIA	718
DB	679	NSTVMSRAEYFNVVDYLLIHGTTADDNVHVFQSAQISKALVNAQVDFQAMWYDQNHGL-S	737
QY	719	STAHOHIYTHMSHFIKQCFSL	739
DB	738	GLSTNHLVTHMTHFLKQCFSL	758

Search completed: February 15, 2006, 13:14:37  
 Job time : 237 secs



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2006, 13:10:58 ; Search time 44 Seconds  
(without alignments)  
1618.191 Million cell updates/sec

Title: US-10-659-055-3  
Perfect score: 4020  
Sequence: 1 ADPGSHHHHSRKTTLT.....AHQHYTHMSHFIRKQFSLP 740

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_80.\*  
1: PIR1.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3933	97.8	766	1 CDHU26	dipeptidyl-peptida
2	3407.5	84.8	792	1 A39914	dipeptidyl-peptida
3	3386	84.2	760	1 S23752	dipeptidyl-peptida
4	1960.5	48.8	759	2 I38593	fibroblast activat
5	1129	28.1	803	2 I68600	dipeptidyl aminope
6	1129	28.1	865	2 I54331	dipeptidyl aminope
7	1106	27.5	803	2 A41793	dipeptidyl aminope
8	950.5	23.6	793	2 T41703	dipeptidyl aminope
9	914	22.7	818	1 A30107	dipeptidyl aminope
10	799.5	19.9	711	2 S66261	X-Pro dipeptidyl-p
11	745.5	18.5	829	2 T19514	hypothetical prote
12	733	18.2	931	2 A49737	dipeptidyl aminope
13	725	18.0	799	2 T25174	hypothetical prote
14	720	17.9	779	2 T25173	hypothetical prote
15	628	15.6	738	2 A87516	dipeptidyl-peptida
16	594.5	14.8	741	2 JC5142	X-Pro dipeptidyl-p
17	590.5	14.7	743	2 T37700	probable dipeptidyl
18	445.5	11.1	931	2 T32919	hypothetical prote
19	334.5	8.3	795	2 T82858	dipeptidyl-peptida
20	272	6.8	657	2 E70025	probable acylamino
21	244.5	6.1	764	2 JC8016	acylaminoacyl-pept
22	226	5.6	622	2 T71174	hypothetical prote
23	221	5.5	591	2 T72474	probable acylamino
24	219	5.4	642	2 C71137	hypothetical prote
25	214	5.3	632	2 E75057	peptidase PAB1418
26	205	5.1	631	2 E75007	probable acylamino
27	205	5.1	683	2 E87495	prolyl oligopeptid
28	191	4.8	709	2 B82580	alanyl dipeptidyl
29	188	4.7	659	2 F72568	probable acylamino

ALIGNMENTS

RESULT 1

CDHU26

dipeptidyl-peptidase IV (EC 3.4.14.5) - human

N;Alternate names: cell surface glycoprotein CD26; thymocyte-activating molecule (THAM)

C;Species: Homo sapiens (man)

C;Date: 31-Dec-1993 #sequence revision 23-Aug-1996 #text change 09-Jul-2004

C;Accession: S24313; B42408; A42408; B61136; S59510; I56154; S59857; S15520

R;Misumi, Y.; Hayashi, Y.; Arakawa, F.; Ikehara, Y.

Biochim. Biophys. Acta 1131, 333-336, 1992

A;Title: Molecular cloning and sequence analysis of human dipeptidyl peptidase IV, a ser

A;Reference number: S24313; MUID:92329551; PMID:1352704

A;Accession: S24313

A;Molecule type: mRNA

A;Residues: 1-6, 'I', 8-766 <MIS>

A;Cross-references: UNIPARC:UPI000016A880; EMBL:X60708; NID:G953335; PIR

R;Darmoul, D.; Lacasa, M.; Baricault, L.; Marguet, D.; Sapin, C.; Trotot, P.; Barbat, P.

J. Biol. Chem. 267, 4824-4833, 1992

A;Title: Dipeptidyl peptidase IV (CD 26) gene expression in enterocyte-like colon cancer

IV mRNA levels during cell differentiation

A;Reference number: A42408; MUID:92165847; PMID:1347043

A;Accession: B42408

A;Molecule type: mRNA

A;Residues: 1-5, 'R', 7-436, 'S', 438-556, 'I', 558-662, 'E', 664-766 <DAR1>

A;Cross-references: UNIPARC:UPI000052ACB; GB:M80536; NID:G181569; PIDN:AAA52308.1; PIR

A;Experimental source: intestine

A;Note: this sequence corresponds with the author's translation

A;Accession: A42408

A;Molecule type: mRNA

A;Residues: 1-5, 'R', 7-436, 'S', 438-556, 'I', 558-662, 'E', 664-711, 'G', 713-766 <DAR2>

A;Cross-references: UNIPARC:UPI0000172A2B; GB:M80536; NID:G181569

A;Note: sequence extracted from NCBI backbone (NCBIN:83986, NCBI:83988); this sequence

R;Gorvel, J.P.; Ferrero, A.; Chambrault, L.; Rigal, A.; Bonicel, J.; Maroux, S.

Gastroenterology 101, 618-625, 1991

A;Title: Expression of sucrose-isomaltase and dipeptidylpeptidase IV in human small int

A;Reference number: A61136; MUID:91317403; PMID:1677636

A;Accession: B61136

A;Molecule type: protein

A;Residues: 1-15, 'X', 17-22 <GOR>

A;Cross-references: UNIPARC:UPI0000172A2C

R;Boehm, S.K.; Gum Jr., J.R.; Erickson, R.H.; Hicks, J.W.; Kim, Y.S.

Biochem. J. 311, 835-843, 1995

A;Title: Human dipeptidyl peptidase IV gene promoter: tissue-specific regulation from a

A;Reference number: S59510; MUID:96067599; PMID:7487939

A;Accession: S59510

A;Molecule type: DNA

A;Residues: 1-31 <BOE>

A;Cross-references: UNIPARC:UPI000016B4A6; GB:S79876; NID:G1195574; PIDN:AA835614.1; PIR

R;Tanaka, T.; Camerini, D.; Seed, B.; Torimoto, Y.; Dang, N.H.; Kameoka, J.; Dahlberg,

J. Immunol. 149, 481-486, 1992

A;Title: Cloning and functional expression of the T cell activation antigen CD26.

A;Reference number: I56154; MUID:92325476; PMID:1352530

A;Accession: I56154

A>Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-436,'S',438-766 <TAN>  
A:Cross-references: UNIPARC:UPI000004F7BF; GB:M74777; NID:gl80082; PIDN:AAA51943.1; PID:  
R.Abbott, C.A.; Baker, E.; Sutherland, G.R.; McCaughan, G.W.  
Immunogenetics 40, 331-338, 1994  
A>Title: Genomic organization, exact localization, and tissue expression of the human CD  
A:Reference number: S59857; MUID:95012454; PMID:7927537  
A:Accession: S59857  
A:Molecule type: DNA  
A:Residues: 1-436,'S',438-766 <ABB>  
A:Cross-references: UNIPARC:UPI000004F7BF; EMBL:U13734  
C:Genetics:  
A:Gene: GDB:DDP4  
A:Cross-references: GDB:125239; OMIM:102720  
A:Map position: 2q24.3-2q24.3  
A:Introns: 2/3; 32/1; 65/1; 95/3; 122/3; 140/2; 164/3; 205/1; 258/3; 296/2; 341/3; 356/3  
C:Superfamily: dipeptidyl-peptidase IV  
C:Keywords: dipeptidyl-peptide hydrolase; glycoprotein; homodimer; proteinase; transmembr  
F1-6/Domain: intracellular #status predicted <INT>  
F17-28/Domain: transmembrane #status predicted <EXT>  
F29-766/Domain: extracellular #status predicted <EXT>  
F:85,92,150,219,229,281,321,520,685/Binding site: carbohydrate (Asn) (covalent) #status  
F:630,708,740/Active site: Ser, Asp, His #status predicted

Query Match	97.8%;	Score 3933;	DB 1;	Length 766;
Best Local Similarity	99.9%;	Pred. No. 6.5e-259;		
Matches 727;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY 13	SRKTYLTLDYKNTYRLKLYSLRWISDHEYLKQENILVFNAYGNSVFLNSTFDEF	72		
DB 39	SRKTYLTLDYKNTYRLKLYSLRWISDHEYLKQENILVFNAYGNSVFLNSTFDEF	98		
QY 73	GHSINDYSISPGQFILLBYNKKWRHSYTSYDIYDLNKKQLITEERIPNNTQVWTWS	132		
DB 99	GHSINDYSISPGQFILLBYNKKWRHSYTSYDIYDLNKKQLITEERIPNNTQVWTWS	158		
QY 133	PVGHKLAYVWNDIYVKLEPNLPSYRITWTGKEDIYNGITDWVVEEVPFSAYSALWWS	192		
DB 159	PVGHKLAYVWNDIYVKLEPNLPSYRITWTGKEDIYNGITDWVVEEVPFSAYSALWWS	218		
QY 193	NGTFLAYAQFNTEVPLIEYSFYSDLSQYKPTVRVPYKAGVNPVTFVFNVDLSLS	252		
DB 219	NGTFLAYAQFNTEVPLIEYSFYSDLSQYKPTVRVPYKAGVNPVTFVFNVDLSLS	278		
QY 253	VTNATSIQITAPASMLIGDHLYLQVTDVATQERISQWLRRIQNSYVMDICDYDESSGRWN	312		
DB 279	VTNATSIQITAPASMLIGDHLYLQVTDVATQERISQWLRRIQNSYVMDICDYDESSGRWN	338		
QY 313	CLVARQHIEMSTGWGRPRPSPHPTLDGNSFYKLIISNEEGVYRHICYFQIDKDCCTFIT	372		
DB 339	CLVARQHIEMSTGWGRPRPSPHPTLDGNSFYKLIISNEEGVYRHICYFQIDKDCCTFIT	398		
QY 373	KGTWEVIGTEALTSYLLIYSNEYKMGPCGRNLYKIQLSDYTKVTCLSCELNPERCQYYS	432		
DB 399	KGTWEVIGTEALTSYLLIYSNEYKMGPCGRNLYKIQLSDYTKVTCLSCELNPERCQYYS	458		
QY 433	VSFSEAKTYQRCSPGLPLTYLTHSSVNDKGLRVLEDNSALDKMLQNVQMSKKLDFII	492		
DB 459	VSFSEAKTYQRCSPGLPLTYLTHSSVNDKGLRVLEDNSALDKMLQNVQMSKKLDFII	518		
QY 493	LNETHFWYQMLPPHFDKSKYPLLLDLYVAGPCSKADTVFRLNWTATYLASTENIIVASF	552		
DB 519	LNETHFWYQMLPPHFDKSKYPLLLDLYVAGPCSKADTVFRLNWTATYLASTENIIVASF	578		
QY 553	DGRGSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGYVTSM	612		
DB 579	DGRGSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGYVTSM	638		
QY 613	VLGSGSGVPCGIAVAPSRWEYSDVYTERVNGLPTPDNDLDHYNSTVMSRAENFKQV	672		
DB 639	VLGSGSGVPCGIAVAPSRWEYSDVYTERVNGLPTPDNDLDHYNSTVMSRAENFKQV	698		

QY 673 EYLLIHGTADDNVHVFQSSAQISKALVDVGVDFQAMWYTDHGHGIASSAHQHIYTHMSHF 732  
DB 699 EYLLIHGTADDNVHVFQSSAQISKALVDVGVDFQAMWYTDHGHGIASSAHQHIYTHMSHF 758  
QY 733 IKQCFSLP 740  
DB 759 IKQCFSLP 766

RESULT 2  
A39914  
dipeptidyl-peptidase IV [EC 3.4.14.5], membrane-bound form precursor - rat  
N:Alternate names: GP110; membrane glycoprotein 110K; OX-61  
N:Contains: dipeptidyl-peptidase IV, soluble form  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 09-Jul-2004  
A:Accession: A39914; A33315; B33315; A60730; A42203; S38949; A31781  
R:Hong, W.; Doyle, D.  
Proc. Natl. Acad. Sci. U.S.A. 84, 7962-7966, 1987  
A>Title: cDNA cloning for a bile canalicular domain-specific membrane glycoprotein of ra  
A:Reference number: A39914; MUID:88068516; PMID:3479775  
A:Accession: A39914  
A:Molecule type: mRNA  
A:Residues: 1-792 <HON>  
A:Cross-references: UNIPROT:P14740; UNIPARC:UPI000017098A; GB:J02997; NID:G204463; PIDN:  
R:Ogata, S.; Miumi, Y.; Ikehara, Y.  
J. Biol. Chem. 264, 3596-3601, 1989  
A>Title: Primary structure of rat liver dipeptidyl peptidase IV deduced from its cDNA an  
A:Reference number: A33315; MUID:89123496; PMID:2563382  
A:Accession: A33315  
A:Molecule type: mRNA  
A:Residues: 1-37,'A',39-182,'I',184-331,'T',333-351,'C',353-393,'V',395-561,'L',563-623,  
A:Cross-references: UNIPARC:UPI0000129841; GB:J04591; NID:Q203973; PIDN:AAA41096.1; PID:  
A>Note: the authors translated the codon GCG for residue 38 as Arg, ACC for residue 332  
A:Accession: B33315  
A:Molecule type: protein  
A:Residues: 1-20;35-54;427-443;505-509;511-520;530-538;593-600;602-608;618-627 <HO2>  
A:Cross-references: UNIPARC:UPI0000172A37  
A36; UNIPARC:UPI0000172A37  
R:McCaughan, G.W.; Wickson, J.E.; Creswick, P.F.; Gorrell, M.D.  
Hepatology 11, 534-544, 1990  
A>Title: Identification of the bile canalicular cell surface molecule GP110 as the ectop  
Quence.  
A:Reference number: A60730; MUID:90228896; PMID:1970322  
A:Accession: A60730  
A:Molecule type: protein  
A:Residues: 28-47,'XX',50-53,55-58 <MCC>  
A:Cross-references: UNIPARC:UPI0000172A38  
R:Ogata, S.; Miumi, Y.; Tsuji, E.; Takami, N.; Oda, K.; Ikehara, Y.  
Biochemistry 31, 2582-2587, 1992  
A>Title: Identification of the active site residues in dipeptidyl peptidase IV by affini  
A:Reference number: A42203; MUID:92190188; PMID:1347701  
A:Accession: A42203  
A:Molecule type: protein  
A:Residues: 'R',625-630,'X',632-648 <OG2>  
A:Cross-references: UNIPARC:UPI0000172A39  
R:Iwaki-Egawa, S.; Watanabe, Y.; Fujimoto, Y.  
Biol. Chem. Hoppe-Seyler 374, 973-975, 1993  
A>Title: N-terminal amino acid sequence of the 60-kDa protein of rat kidney dipeptidyl p  
A:Reference number: S38949; MUID:94128239; PMID:7905271  
A:Accession: S38949  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 281-302 <IWA>  
A:Cross-references: UNIPARC:UPI0000172A3A  
R:Hong, W.; Doyle, D.  
J. Biol. Chem. 263, 16892-16898, 1988  
A>Title: Membrane orientation of rat gp110 as studied by in vitro translation.  
A:Reference number: A31781; MUID:89034185; PMID:3182821  
A:Accession: A31781  
A:Molecule type: mRNA  
A:Residues: 1-40 <HO3>  
A:Cross-references: UNIPARC:UPI0000172A3B

C:Comment: This protein is localized to the bile canalculus, which is the apical domain  
C:Superfamily: dipeptidyl-peptidase IV  
C:Keywords: dipeptidyl-peptidase hydrolase; glycoprotein; homodimer; liver; serine protease  
F:1-792/Product: dipeptidyl-peptidase, membrane-bound form #status experimental <MATW>  
F:1-28/Domain: signal sequence #link MATS #status experimental <SIG>  
F:1-6/Domain: intracellular #status predicted <INT>  
F:17-28/Domain: transmembrane #status predicted <TMN>  
F:29-792/Domain: extracellular #status predicted <EXT>  
F:35-792/Product: dipeptidyl-peptidase, soluble form #status experimental <MATS>  
F:83,90,148,217,227,319,521,686/Binding site: carbohydrate (Aen) (covalent) #status pred  
F:631/Active site: Ser #status experimental  
F:709,741/Active site: Asp, His #status predicted

Query Match 84.8%; Score 3407.5; DB 1; Length 792;  
Best Local Similarity 84.5%; Pred. No. 3.2e-223;  
Matches 617; Conservative 54; Mismatches 56; Indels 3; Gaps 1;  
Qy 13 SRKTYTLTDYLVKNTYRLKLSLRWISDHEVLYKQENILVFNAYGNSVPLENSTPDEF 72  
Db 37 SRRTYTLADYLVKNTYRLKLSLRWISDHEVLYKQENILVFNAYGNSVPLENSTPDEF 96  
Qy 73 GHSINDYSISPDGQFILLVYVYKQWRHSYTSYDIYDLNKRQLITERRIPNNTQWTVWS 132  
Db 97 GDSISDYSVSPDRFLVLLVYVYKQWRHSYTSYDIYDLNKRQLITERRIPNNTQWTVWS 156  
Qy 133 PVGHKLAVVWNNDIYVKEPNLPSYRIWTGKEDIYNGITDWVYEEVFSAYGALMWSP 192  
Db 157 QEGHKLAVVWNNDIYVKEPNLPSYRIWTGKEDIYNGITDWVYEEVFSAYGALMWSP 216  
Qy 193 NGTFLAYAQFNDTVPLIEYFYSDLSQYPTKTVVPYKAGVNPVTKFVNTDLSLS 252  
Db 217 NGTFLAYAQFNDTVPLIEYFYSDLSQYPTKTVVPYKAGVNPVTKFVNTDLSLS 276  
Qy 253 VTNATSIQTAPASMLGDHYLCVDTWATQBRISLOWLRRIQNTSVMDICDYDESSGRWN 312  
Db 277 TTTTIPMQITAPASVTTGDHYLCVDTWATQBRISLOWLRRIQNTSVMDICDYDESSGRWN 336  
Qy 313 CLVARQHIEMSTGTVGWRFRPSPHFTLDGNSFYKISNIEGYSYHICVFQIDKK---DCT 369  
Db 337 CPTQEHLETSATGTVGWRFRPSPHFTLDGNSFYKISNIEGYSYHICVFQIDKK---DCT 396  
Qy 370 FITKGTWEVIGIEALTSYLYISNEYKMPGGRNLYKIQLSDYTKVTCLSCELNPERCQ 429  
Db 397 FITKGTWEVIGIEALTSYLYISNEYKMPGGRNLYKIQLSDYTKVTCLSCELNPERCQ 456  
Qy 430 YYSVSFSEAKYQYLRCSGPGPLPLYTLHSSVNDKGLVLENSALDKMLQVQMPKSLD 489  
Db 457 YYSVLSSEAKYQYLRCSGPGPLPLYTLHSSVNDKGLVLENSALDKMLQVQMPKSLD 516  
Qy 490 FIILNETKFWYQMLPLPHFDKSKYPLLLDYVYAGPCQKADTVFRLNWTYLASTENIIV 549  
Db 517 FIVLNETKFWYQMLPLPHFDKSKYPLLLDYVYAGPCQKADTVFRLNWTYLASTENIIV 576  
Qy 550 ASPDGRSGYQGDKIMHAINRRLGTFFVEDQIEAARQFSKMGFVNDKRIATIGWSYGYV 609  
Db 577 ASPDGRSGYQGDKIMHAINRRLGTFFVEDQIEAARQFSKMGFVNDKRIATIGWSYGYV 636  
Qy 610 TSMVLGSGSGYKFCGIAVAPSRVEYVDSVYTERYMGILPTPEDNLDHYRNSVMSRAENF 669  
Db 637 TSMVLGSGSGYKFCGIAVAPSRVEYVDSVYTERYMGILPTPEDNLDHYRNSVMSRAENF 696  
Qy 670 KQVEYLLHGTADDNVHFQQAQISKALVDVGVDFQAWWYTDDEHGIASSTAHQHIIYTHM 729  
Db 697 KQVEYLLHGTADDNVHFQQAQISKALVDVGVDFQAWWYTDDEHGIASSTAHQHIIYTHM 756  
Qy 730 SHFIKQCFSL 739  
Db 757 SHFIKQCFSL 766

RESULT 3  
S23752

dipeptidyl-peptidase IV (EC 3.4.14.15) alpha chain - mouse  
N:Alternate names: CD26 alpha subunit; THAM alpha subunit  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: S23752; A46465; A56030  
R:Marguet, D.; Bernard, A.M.; Vivier, I.; Darmoul, D.; Naquet, P.; Pierres, M.  
J. Biol. Chem. 267, 2200-2208, 1992  
A:Title: cDNA cloning for mouse thymocyte-activating molecule. A multifunctional ecto-  
A:Reference number: S23752; MUID:92129288; PMID:1370813  
A:Accession: S23752  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-760 <MAR>  
R:Cross-references: UNIPARC:UPI0000172A2D; EMBL:X58384  
R:Vivier, I.; Marguet, D.; Naquet, P.; Bonicel, J.; Black, D.; Li, C.X.; Bernard, A.M.  
J. Immunol. 147, 447-454, 1991  
A:Title: Evidence that thymocyte-activating molecule is mouse CD26 (dipeptidyl peptida-  
A:Reference number: A46465; MUID:91302787; PMID:1712807  
A:Accession: A46465  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-20 <VIV>  
R:Cross-references: UNIPARC:UPI0000172A2E  
A:Experimental source: M14.T thymoma cells, Swiss nu/nu  
A:Note: sequence extracted from NCBI backbone (NCBIP:42236)  
R:Bernard, A.M.; Mattei, M.G.; Pierres, M.; Marguet, D.  
Biochemistry 33, 15204-15214, 1994  
A:Title: Structure of the mouse dipeptidyl peptidase IV (CD26) gene.  
A:Reference number: A56030; MUID:95092780; PMID:7999781  
A:Accession: A56030  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 746-760 <BER>  
R:Cross-references: UNIPARC:UPI0000172A2F; GB:U12620  
C:Genetics:  
A:Gene: CD26  
C:Superfamily: dipeptidyl-peptidase IV  
C:Keywords: dipeptidyl-peptidase hydrolase; glycoprotein; transmembrane protein  
F:213,223,315,514,675/Binding site: carbohydrate (Aen) (covalent) #status predicted  
F:624,702,734/Active site: Ser, Asp, His #status predicted  
Query Match 84.2%; Score 3386; DB 1; Length 760;  
Best Local Similarity 84.6%; Pred. No. 8.6e-222;  
Matches 615; Conservative 57; Mismatches 51; Indels 4; Gaps 1;  
Qy 13 SRKTYTLTDYLVKNTYRLKLSLRWISDHEVLYKQENILVFNAYGNSVPLENSTPDEF 72  
Db 37 SRRTYSLADYLVKNTYRLKLSLRWISDHEVLYKQENILVFNAYGNSVPLENSTPDEF 96  
Qy 73 GHSINDYSISPDGQFILLVYVYKQWRHSYTSYDIYDLNKRQLITERRIPNNTQWTVWS 132  
Db 97 GY----HSVSPDRFLVLLVYVYKQWRHSYTSYDIYDLNKRQLITERRIPNNTQWTVWS 152  
Qy 133 PVGHKLAVVWNNDIYVKEPNLPSYRIWTGKEDIYNGITDWVYEEVFSAYGALMWSP 192  
Db 153 PEGHKLAVVWNNDIYVKEPNLPSYRIWTGKEDIYNGITDWVYEEVFSAYGALMWSP 212  
Qy 193 NGTFLAYAQFNDTVPLIEYFYSDLSQYPTKTVVPYKAGVNPVTKFVNTDLSLS 252  
Db 213 NNTFLAYAQFNDTVPLIEYFYSDLSQYPTKTVVPYKAGVNPVTKFVNTDLSLS 272  
Qy 253 VTNATSIQTAPASMLGDHYLCVDTWATQBRISLOWLRRIQNTSVMDICDYDESSGRWN 312  
Db 273 SSSNAPLQIPAPASVARGDHYLCVDTWATQBRISLOWLRRIQNTSVMDICDYDESSGRWN 332  
Qy 313 CLVARQHIEMSTGTVGWRFRPSPHFTLDGNSFYKISNIEGYSYHICVFQIDKKDCTFIT 372  
Db 333 CPSEQQHVENSTGTVGWRFRPSPHFTLDGNSFYKISNIEGYSYHICVFQIDKKDCTFIT 392  
Qy 373 KGTWEVIGIEALTSYLYISNEYKMPGGRNLYKIQLSDYTKVTCLSCELNPERCQYS 432  
Db 393 KGAWEVISIALTSYLYISNEYKMPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYA 452

```

QY 433 VSFSEAKYQYLCRSGPLPLTLHSSVNDKGLRVLEDSALDKMLQNVQMPSSKLDPLII 492
Db 453 VSFSEAKYQYLCRSGPLPLTLHSSVNDKGLRVLEDSALDKMLQNVQMPSSKLDPLII 512
QY 493 LNETKFWYQMLPPHFDKSKYPLLLDYYVAGPCSKADTVFRLNWTYLASTENIIVASF 552
Db 513 LNETKFWYQMLPPHFDKSKYPLLLDYYVAGPCSKADTVFRLNWTYLASTENIIVASF 572
QY 553 DGRSGYQGDQKIMHAINRRLGTFFVEDQIEAARQPSKMGFVNDKRIAIWGSYGGYVTSM 612
Db 573 DGRSGYQGDQKIMHAINRRLGTFFVEDQIEAARQPSKMGFVNDKRIAIWGSYGGYVTSM 632
QY 613 VLGSQGVKFCGIAVAPVSRWEYDVSVYTERYMGLPDPEDNLDHYRNSVMSRAENPKQV 672
Db 633 VLGSQGVKFCGIAVAPVSRWEYDVSVYTERYMGLPDPEDNLDHYRNSVMSRAENPKQV 692
QY 673 EYLLHGTADDNVHFOQSAQISKALVDGVDFQAMWYTDDEHGIASTAHQHIYTHMSHF 732
Db 693 EYLLHGTADDNVHFOQSAQISKALVDGVDFQAMWYTDDEHGIASTAHQHIYTHMSHF 752
QY 733 IQQCFSL 739
Db 753 LQQCFSL 759

RESULT 4
138593
fibroblast activation protein-alpha - human
N:Alternate names: FAP-alpha
C:Species: Homo sapiens (man)
C>Date: 09-Mar-1996 #sequence_revision 09-Mar-1996 #text_change 29-Aug-1997
C:Accession: I38593
R:Scanlan, M.J.; Raj, B.; Calvo, B.; Garin-Chesa, P.; Sanz-Moncasi, M.P.; Healey, J.; Ol
Proc. Natl. Acad. Sci. U.S.A. 91, 5657-5661, 1994
A:Title: Molecular cloning of fibroblast activation protein alpha, a member of the sexin
A:Reference number: I38593; MUID:94561645; PMID:7911242
A:Accession: I38593
A>Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-759 <RES>
A:Cross-references: UNIPARC:UPI000175B7D; EMBL:U09278; NID:g507749; PID:g507750
C:Genetics:
A:Gene: GDB:PAP
A:Cross-references: GDB:374184; OMIM:600403
A:Map position: 2q23-2q23
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: transmembrane protein

Query Match 48.8%; Score 1960.5; DB 2; Length 759;
Best Local Similarity 48.5%; Pred. No. 4.8e-125; Indels 35; Gaps 11;
Matches 364; Conservative 138; Mismatches 214;

QY 3 PGSSHHHHHSRKYTLTYLKNYLRKLYSLRWISDHELYLKQ-ENNILVFNAYEGNSS 61
Db 28 PSRVHSENTWRLTKDLNGTFSYKTFPFNWSIGQVEYLHQSADNNVLYNIETGQSY 87
QY 62 VFLENSTDFPFGHSIN--DYSISPGQFTILLYNVVQWRHSYTSYDIYDLNKKQLITE 119
Db 88 TILSNRTM---KSNVNASNYGLSPDRQFVYLSDSYSLKWRYSYATYTYIDLSNGEFVRG 143
QY 120 ERIPNNQVTVSPVCHKLAYVNNNDIYVKIEPNLPSYRIWTGKEDIYNGITDWYEE 179
Db 144 NELPRFIQLCWSPVGSKLAYVYQNNIYLKQRPDPPFQITFGRENKIFNGIPDWYEE 203
QY 180 EVFSAYSALWSPNGTFLAYAFNDTEVPLIYSFYSDESLOQPKTVRPVYPKAGAVNPT 239
Db 204 EMLPTKYLWSPNGKFLAYAFNDKDIPIVAYSYYGDE--QYPTINIPYKAGAKNPV 261
QY 240 VKFPVNTOSLSVNTATSIQITAPASMLIGHYLCVDTWATQERISLOWLRRIQNYSYM 299
Db 262 VRIFIIDTTPAYVGPQ---EVPVPAMIASSDYFSLWLTWTDERVCLQWLKRVQNSYL 318
QY 300 DICVDYDESSGRNCLVARQHIEMTGTGWRFRPSPHPTLDGNSPKYIISNEEYRHIC 359

```

```

Db 319 SICDFREDQWDCPKTQEHIEESRTWAGGFFVSRPVSFSDAIYVYKIFSCKDGYKHH 378
QY 360 YFQIDKDCCTPITKGTWEVIGIEALTSYLYISNEYKMGPCGRNLYKIQLSDY-TKVT 418
Db 379 YIKDTVENAIIQTSGKEAINIFRVTDLSLFYSSNEFEYFGRRNIYRISIGSPPSKKC 438
QY 419 LSCLEINPERCOVYSVSFSKEAKYTYOLRCGCLPLTYTLHSSVNDKGLRVLEDSALDKWL 478
Db 439 VTCHLRKERCQYTYTASFDYAKYVALVCYGPPIPLSTLHDGRTDQEIKLEENKELENAL 498
QY 479 QNVQMPSSKLDPLIIINETKFWYQMLPPHFDKSKYPLLLDYYVAGPCSKADTVFRLNWA 538
Db 499 KNIQLPKEIEKKLEVEDEITLWYKMLLPQFDRSKYPLLIQYVGPCSVSRVSFAVAMI 558
QY 539 TYLASTENIIVASFGRSGYQGDQKIMHAINRRLGTFFVEDQIEAARQPSKMGFVNDKRI 598
Db 559 SYLASKEGWVIALVDGRGTAFQGDKLLYAVYRKLGVYEVEDQITAVRKFIEMGFIDEKRI 618
QY 599 AIWGSYSG-GYVT-----SMVLGSSGVKFCGIAVAPVSRWEYDVSVYTERYMGLP 648
Db 619 AIWGSYSEIRFITGPCIWNWSFQMYSSGSL-QGLRLRVCLHR-----EIHGSP 667
QY 649 TPEDNLDHYRNSVMSRAENPKQVEYLLIHGTADDNVHFOQSAQISKALVDGVDFQAMW 708
Db 668 NKDDNLEHYKNSVWMAEYFRNVYLLIHGTADDNVHFOQSAQISKALVNAQVDFQAMW 727
QY 709 YTDHGHGIASTAHQHIYTHMSHFPIKQCFSL 739
Db 728 YSDQNHGL-SGLSTNHLTYTHMTFLKQCFSL 757

RESULT 5
168600
dipeptidyl aminopeptidase like protein - human
C:Species: Homo sapiens (man)
C>Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000
C:Accession: I68600
R:Yokotani, N.; Doi, K.; Wenthhold, R.J.; Wada, K.
Hum. Mol. Genet. 2, 1037-1039, 1993
A:Title: Non-conservation of a catalytic residue in a dipeptidyl aminopeptidase IV-relat
A:Reference number: I54331; MUID:93372805; PMID:8103397
A:Accession: I68600
A>Status: Preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-803 <RES>
A:Cross-references: UNIPARC:UPI000016A7F9; GB:M96860; NID:g306707; PIDN:AAA35761.1; PID:
C:Superfamily: dipeptidyl-peptidase IV

Query Match 28.1%; Score 1129; DB 2; Length 803;
Best Local Similarity 32.7%; Pred. No. 1.4e-68;
Matches 244; Conservative 155; Mismatches 303; Indels 44; Gaps 20;

QY 14 RRTYTLTYLKNYLRKLYSLRWISDHELYLKQ-ENNILVFNAYEGNSSVFLNSTDDEF 72
Db 66 KCKVTVEDELPSDDPKIHDPKAWISDTEFIYREQGTVRLWNVETNTSVLIEGKKIESL 125
QY 73 GHSINDYISDPGQFTILLYNVVQWRHSYTSYDIYDL--NKRQLITEIRIPN-NQOV 129
Db 126 -RAIR-YEISPDREYALFSYVNEPIYQHSYTYGTYVLSKIPHGDPOSQDLPPEVSNAKLQYA 183
QY 130 TWSPVGHKLAYVNNNDIYVKIEPNLPSYRIWTGKEDIYNGITDWYEEEFVSAYSALW 189
Db 184 GWPGKQQLIFIFENNIYCAHVQKQARVSVTGKEGYINGLSWLYEEELKTHIAHW 243
QY 190 MSPNGTFLAYAFNDTEVPLIYSFYSDESLOQPKTVRPVYPKAGAVNPTVFFVNTDS 249
Db 244 MSPDGTFLAYAINDSRPVIMELPYTGSI--YPTVKPVHYPKAGSENPSISLHVI---G 298
QY 250 LSSVNTATSIQITAPASMLIGHYLCVDTWATQERISLOWLRRIQNYSYMDCDDESSG 309
Db 299 LNPPTH--DLEMPMPDDPRMREYITVMKWTATSTKVAVTMLNRAQNVSLTLC--DATG 354

```

Qy	310	RWNCIVARQHIEMSTTGWVRPPRPEPHPTLDGNSFYKIIISNEBGYRHCYP-----	361
Db	355	-----VCTKKH-EDESEAWLHR--ONEEPVFSKGRKFFPFIIRAIPQGGRGKFIHITVSSSQP	408
Qy	362	QIDKKDCFTFKGTWEVIGIEAL--TSDVLYIYSNEYKMGPGCRNLYKIQLSDTYKVTCL	419
Db	409	NSSNDNIOSITSGDWDVTKLAYDEGKNKIYFLSTE--DLPRRQLIYSANTEGNFRQCL	466
Qy	420	SCELNPERCQYYSVSKEAKYQIURCSGPGUPLYTLHSSVNDKGLRVLIEDNSALDKMLQ	479
Db	467	SCDL--VENCITYFSASFHSMDPFLKCKECPGVPMVTHNTTDKCKMFDLETNEHVKKAIN	525
Qy	480	NVQMSKCLDRIILNETKFTWQMIILPPHPDKSKCYPLLLDVYAGPCSKQADTVFLRWAT	539
Db	526	DRQMPKVEYRIEIDBYNLPQILKPAFTDTHYPLLLVDGTPGSQSVASKEFVSWET	585
Qy	540	YLASTENIIVASFDRGSGYQGDKIMHAINRRLGTFEVEDQIEAARQPSKMGFVDMKRIA	599
Db	586	VWVSSHGAVVVKCDGRGSGFGQTKLLHVEVRRLGLLEBKQMEAVRTMLKEQYIDRTRVA	645
Qy	600	IWHSYSGYVYTSWL-----GSSGG--VPKCGIAVAPVSRWEHYDSVTERYMGUPTPEDNLD	655
Db	646	VFGKDYGYLZTYILPAKGENGQGTFTCGSALSPTDFKLYASAFSERYLGL-----HGJD	701
Qy	656	H--YRNSNTWMAEKNFKOVBYLLIHGTADDNVHFQSOAQISKALVDVGVDFQAMWYTDSD	713
Db	702	NEAYEMTKVAIRVSALSEQOFLIHPTADEKLIHFQHTAELITQILNGKANYSLQIYPDES	761
Qy	714	HGIASSTAQHYYTHMSHFIKQCPSL	739
Db	762	HYFTSSSLKOHLYRSIINFFVECFRI	787

RESULT 6  
I54331  
dipeptidyl aminopeptidase like protein - human  
C/Species: Homo sapiens (man)  
C/Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
C/Accession: I54331  
R/Yokotani, N.; Doi, K.; Wenthold, R.J.; Wada, K.  
Hum. Mol. Genet. 2, 1037-1039, 1993  
A/Title: Non-conservation of a catalytic residue in a dipeptidyl aminopeptidase IV-related  
A/Reference number: I54331; MUID:93372805; PMID:8103397  
A/Accession: I54331  
A/Status: preliminary; translated from GB/EMBL/DBDJ  
A/Molecule type: mRNA  
A/Residues: 1-865 <RES>  
A/Cross-references: UNIPROT:P42658; UNIPARC:UPI0000047378; GB:M96859; NID:g306705; PIDN:  
C/Superfamily: dipeptidyl-peptidase IV

Query Match	28.1%	Score 1129;	DB 2;	Length 865;
Best Local Similarity	32.7%	Pred. No. 1.6e-68;		
Matches	244;	Conservative 155;	Mismatches 303;	Indels 44;
				Gaps 20
Qy	14	RKTYTLDTYKNTVRLKLYLSRWSDHLYLK-QENNILVFNVAEYGNSSVFLNSTFDEF	72	
Db	128	KKKYTVEDLFSDFPKIHDPKAWISDTEFIREQKGTVRLMNVETNTSTVLIEGKKIESL	187	
Qy	73	GHSINDYSISPDGQFILLVYNNVKWRSSYTSAYDIYDL--NKROLITTEERIPN-NTQWV	129	
Db	188	-RAIR-YEISDPREPLAFSINVEPIYQHSYTGYYVLISKIPHGDPSLODPPEVSNAKLQYA	245	
Qy	130	TWSPVGHKLAVWNNDIYVKLEPNLPSRYITWGTGKBDIIYNGITDQWVEEYRFSAYSALW	189	
Db	246	GWGPKGQOLIFIPENIYYCAHVQKQAIRVSVSTGKEGVYNGLSDWLVEEELKTHIAHW	305	
Qy	190	WSPNGTFLAYAQFNDTEVPLIEYSYSDSESLQYKPTVRVPYPKAGAVNPATKFFVYNTDS	249	
Db	306	WSPDGTRLAYAINDSRVPIMELPYTGTGI--YPTVKVPHYPKAGSENPISLHVI---G	360	
Qy	250	LSSVNTNATSIQITAPASMLIGHYLCDVWTAQERISLQWRRIQNYQVMDICDYDESSG	309	
Db	361	LANGPTH--DLEMPDPDDPRMEYITVMVKWATSTKVAVTMLNRAQNSVLTLC--DATTG	416	

[illegible]

## RESULT 7

R41793  
dipeptidyl aminopeptidase-like protein (EC 3.4.14.-) short form, DPPX-S - bovin  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A41793  
R:Wada, K.; Yokotani, N.; Hunter, C.; Doi, K.; Wenthold, R.J.; Shimaseaki, S.  
Proc. Natl. Acad. Sci. U.S.A. 89, 197-201, 1992  
A:Title: Differential expression of two distinct forms of mRNA encoding member  
A:Reference number: A41793; MUID:92108018; PMID:1729689  
A:Accession: A41793  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-803 <WAD>  
A:Cross-references: UNIPROT:P42659; UNIPARC:UPI000002A93C; GB:M76429; NID:g408  
A:Note: sequence extracted from NCBI backbone (NCBI:P75138)

C;Keywords: dipeptidylpeptide hydrolase; glycoprotein  
P:257.342/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	27.5%	Score 1106;	DB 2;	Length 803;
Best Local Similarity	32.4%	Pred. No. 5.1e-67;		
Matches 242;	Conservative 158;	Mismatches 300;	Indels 48;	Gaps 22

Qy	14	RKVTYLDYLKNTVRLKLYLSURWISDHBYLYK-QBNNTLVFNABYGNSSVFIENSTFDFB 72
Db	66	KKKVTVDELSEDPKIHDPPEAKWISDKFEIYREQKGSVILRNVTNTSVLTJEGKKIESL 125
Qy	73	GHSINDYSISDGGFILLEVYNNVQWRHSYTSASDIYDL--NKRQLITTEERIPN-NTQWV 129
Db	126	-RAIR-YEISPDREYALFSYNVEFIQHSYTYGYVLSKIPHGDPQSLDPPEVSNAKLQYA 183
Qy	130	TWSPVGHKLAVWNNDIYVKEPNLPSVRIWTGKEDIYINGITDPAWTEEEVFSAYSALW 189
Db	184	GMGPKGQQLPFIENNYYCAHVQKQAIRVVSTGKGVYINGLSLDWLYEEELTKTHIAHW 243
Qy	190	WSPNGTFLAYAQFNDTEVPLLEYGFYSDESLOYPKTVRVYPKAGAVNPTPKFVVNTDS 249
Db	244	WSPDGTFLAVATINDSRVPVNNELPTTGS--VYPTAKPVHYPKACENPSISLHVI---G 298

```
QY 250 LSSVTNATSIQTAPASMLIGDHYLCDVTWATOERISLQWRRIQNYSDICDYDESSG 309
DB 299 LNPGRH--DLETPPDDPRMRREYIITWVKWATSTKVAVNWLGRAQNVSLTLC--DAITG 354
QY 310 RWNCLVARQHIEMSTTGWGRFRPSPHPHTLDCNSFYKIIISNEEG----YRHICYF---- 361
DB 355 ----VCTKQH-EDESAHLR-QNEPVFSKGRKFFFYRAIPQGGQCKFYHITVSSSQP 408
QY 362 QIDKDCCTFITKGTWEVIGIALTSD----YLYYISNEYKMPGGRNLKYIQLSDYTKVT 417
DB 409 NSSNDNIQISITGDMVDVTKI--LSYDEKRSQIYFLSTE--DLPRRRQLYSASTVGSFNRQ 464
QY 418 CLSCELNPERCQYYSVSFSEAKYQYLRCSGCLPLYTLHSSVNDKGLRVLSDNSALDKW 477
DB 465 CLSCDL-VNCTYFSASFAGDFFLLKCEGPGVPTVSVHNTTDDKCKMPDLETNEHVQKA 523
QY 478 LQNVOMPSKKLDFIILNETKFWQMLPPLPHFDKSKKYPILLDLYVAGPCQKADTVFRLNW 537
DB 524 ISDRQMPKYEYRKIEITDDYNLPILQILKPATFTDTAHYPLLLVVDGTPGQSQAERKFAVTW 583
QY 538 ATYLASTENIIVASFGRSGGQGDKIMHAINRRLTGTFVEVDQIEAARFQSKMGFVNDKR 597
DB 584 ETVMVSSHGAUVKCDGRSGSGFQTRLLHEVRRRLGSLSEKQDQMEAVRVMLEKPYIDKTR 643
QY 598 IALWGSYGGVYTSMTL---GSGSG-VFKCGIAPVPSWSEYDYSVYTERYMGCLPTPEDN 653
DB 644 VAVFGKDYGGYLYSTYLLPAKGQGAQPVFCGSALSPTITFKUYASAFSRYLGL-----HG 699
QY 654 LDH--YRNTVMSRAENFKQVEYLLHTGTADONVHFQOSAQSIALVDVGVDFQAMWYTD 711
DB 700 LDNRVEMAKVAHRVSAISGQQFLVIHATADEKIHQHTAELITQLIKKANYSLQIYPD 759
QY 712 EDHGIASSTAHOHYIYTHMSHFTKQCFSL 739
DB 760 ESHYFSSAALQOHLHRSILGFFVECPRI 787

RESULT 8
T41703
dipeptidyl aminopeptidase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T41703
R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1998
A:Reference number: Z22011
A:Accession: T41703
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-793 <MUR>
A:Cross-references: UNIPROT:O14073; UNIPARC:UPI000013AB95; EMBL:AL031180; PIDN:CAA20138.
A:Experimental source: strain 972h-; cosmid c2E11 -chimeric
C:Genetics:
A:Gene: SPAC2E11.08
A:Superfamily: dipeptidyl-peptidase IV

Query Match 23.6%; Score 950.5; DB 2; Length 793;
Best Local Similarity 30.5%; Pred. No. 1.8e-56;
Matches 237; Conservative 141; Mismatches 289; Indels 109; Gaps 25;

QY 14 RKTYYTLVDLYNTYRLKLSLRWISDH----EVLVYKQENNILVFNAYGNSSVF-LENST 68
DB 52 RDSLSLDDIVLQYKPSYQVNWIDSQGLKDTPLVKYGLDINIQDPYLNKTLFVSVDLV 111
QY 69 EDEFSGHSINDYSTSPDQFTLLEYNVVKWRHSHYSTASYDIYDLNKLQLITEIRPN---- 124
DB 112 INGIQLDYDYSISLSPFAKTVSVNKSQRWRHSSFAQYLYLN-----TETKDVNMLGQ 164
QY 125 -NTQW----VTWSPVGHKLAIVWNNDIYV-KIEPNLPSYRITWTGKEDIYNGITDWYVE 178
DB 165 DNEHWTISLAESPTGHLFSFVYNNDLVYRKDGNV--QRLTYDGTVD-VFNGLTDWIYE 221
```

```
QY 179 BEVESAYSALWWSPNGTFLVAQAFNDTEVPLIEYSFYSGDE-----SLQYPKTVRVPPYK 232
DB 222 EEVLSPTTIWNSPDSKIAFLKLNESIEIPYHYPLTYAELDPSLPEDFYNDKMAIKYK 281
QY 233 AGAVNPTVKFFVYVNTDSLSSVTNATSIQITAPASMLIGDHYLCDVTWATQERISIQWLRR 292
DB 282 PCGNPNPVSFLV--ADLNSNASSNFSLWHNEP----LAEPVVQVNLVWNTSSVLVQFTNR 335
QY 293 IONYSVMDICDYDESSGRNCLVARQHIEMSTTGWGRFRPSPHPHTLD-----GNS 344
DB 336 NSTCITABRLDTELKS-----IHTVKYTECLEBEGY-EVQQAQKMPFLNNSLVWENWSDG 388
QY 345 FYKIIISNEEGYRHICYFQIDKDCCTFITKGTWEVIG-----IEALTS 387
DB 389 YFDIILALDD-YNHLAIFIPNGSSPYILTSGAWDVTDPHIHIDGDFGNVYFLATLKDSTER 447
QY 388 YLYYISNEYKMPGGRNLKYIQLSDYTKVTCLSCELNPERCQYYSVSFSEAKYIQLRCS 447
DB 448 HLYYYS-----LDTLIYGITD-----NGEDEGYYSTSPFGDFVLYNYH 488
QY 448 GGLPLYTLHSSVNDKGLRVLSDNSALDKMLQNVOMPSKKLDFIILNETKFWQMLPLPH 507
DB 489 GPDVPWQELRSTKOKDYCLSLSTNSRLKQOLSSITLPSVEYKGLTFNDTTFNFMERRPRN 548
QY 508 FDKSKKYPLLLDLYVAGPCQKADTVFRLNWTATYLASTEN--IIVASFQDGRSGSGYQGDKIM 565
DB 549 FDNVKKYPVLFPAYGCGSQQVAKLFRVDVFNQYLAHSDPFEFIVLTDGRTGTFNGNAPR 608
QY 566 HAINRRLTGTFVEVDQIEAARFQSKMGFVNDKRIALWGSYGGVYTSMTVLGSGSGVFKCGI 625
DB 609 YSVSRHLGEWESYDQGAQKFWADLPFVDENHVGWISGYSGLYTLKTLT-ODVFSYGM 667
QY 636 AVAPVSRWEYDYSVYTERYMGCLPTPEDNLDHYRNTVMSRAENFKQVE-YLLIHGTADDN 684
DB 668 AVAPVTDWRLYDVSVYTERYMDL--PQYNKEGYKNSQIHD-YBKQQLKRFVFAHGFGDDN 724
QY 685 VHFQOSAQISKAL-----VDVGVDVFMQAMWYTDDEHGIASSTAHOHYIYTHMSHFI 733
DB 725 VHFQSHMLMDGLNLANCYNDMAV-----FPDSAHSISYHNASLSIYHRUSEWI 774

RESULT 9
A30107
dipeptidyl aminopeptidase B (EC 3.4.14.-) - yeast (Saccharomyces cerevisiae)
N:Alternate names: protease YHR028c
C:Species: Saccharomyces cerevisiae
C>Date: 07-Jun-1990 #sequence_revision 30-May-1997 #text_change 09-Jul-2004
C:Accession: S46780; A30107
R:Du, Z.
submitted to the EMBL Data Library, June 1994
A:Description: The sequence of S. cerevisiae cosmid 8082.
A:Reference number: S46773
A:Accession: S46780
A:Molecule type: DNA
A:Residues: 1-818 <DUZ>
A:Cross-references: UNIPROT:P18962; UNIPARC:UPI0000031A5F; EMBL:U10399; NID:g500689; PFI
R:Roberts, C.J.; Pohlhig, G.; Rothman, J.H.; Stevens, T.H.
J. Cell Biol. 108, 1363-1373, 1989
A:Title: Structure, biosynthesis, and localization of dipeptidyl aminopeptidase B, an in
A:Reference number: A30107; MUID:89174971; PMID:2647766
A:Accession: A30107
A:Molecule type: DNA
A:Residues: 1-82, 'H', 84-124, 'N', 126-181, 'LRRLET', 189-199, 'N', 201-365, 'DFKRGKERKE', 376-5
A:Cross-references: UNIPARC:UPI0000172A3C; EMBL:X15484
A>Note: the authors translated the codon ACC for residue 572 as Asn
C:Genetics:
A:Gene: SGD:DAP2; STE13; MIPS:YHR028c
A:Cross-references: SGD:S0001070; MIPS:YHR028c
A:Map position: 8R
C:Superfamily: dipeptidyl-peptidase IV
C:Keywords: dipeptidylpeptide hydrolase; glycoprotein; transmembrane protein; yeast vacu
F:30-45/Domain: transmembrane #status predicted <TMM>
```



P:63,79,110,139,392,421/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 22.7%; Score 914; DB 1; Length 818;  
 Best Local Similarity 32.3%; Pred. No. 5,7e-54;  
 Matches 258; Conservative 119; Mismatches 314; Indels 108; Gaps 29;

QY 11 HHS-----RKTTLTDYK-----NTYRLKLSLRWISHEVLYKQNNIL--- 51  
 DB 50 HHSNTPDYQEPNSNTDNGKLVSVSVVNNTPQPKYHELOWISDN---KTESNDLGILY 105  
 QY 52 -----VFNAEYGN--SVLENSTDEFGHSINDYSI--SPGQPTILLEYVVKWR 99  
 DB 106 VTFMDSYVKSVDSDNSVLLKGTFIHNGQNLTVESITAPSKLLKRTNSQWR 165  
 QY 100 HSYTASYDIYDLNKRQLITEERIPNNTQVWSPVGHKLAYVNNNDIYVKIEPNLPSYRI 159  
 DB 166 HSTFGSYFYDKSSSF---BEIGNEVALATWSPNSNDIAYVODNNIYISAIKKTIRA 222  
 QY 160 TWTKGEDIYNGITDWTVEEVFSAYSAWSPNGTFLAYAQFNDTEVP--LIEYSFYSD 217  
 DB 223 VTNDGSSFLFNGKPDWVYEEVEFDDKAAMSPGTYLAFKIDSEVGEFTIPY-VYQD 281  
 QY 218 ESLOYPKTVRVPYKAGAVNTPKFPVNTDSSLVVTNATSQITAPASMLIGDHYLVDV 277  
 DB 282 EKDIYPERWSIKPKSGTNPMAELWVY-----SMKDGTSFHPHISGNKKDGLLITEV 335  
 QY 278 TWATOERISLOWLRIONYSVMDICDYDESSGRWNCVLVAROHIEHMTTGW-----VGRFR 332  
 DB 336 TWGNGNVLVKTDRSSDILTVFLDITAKTSN-----VVRN--ESSNGGWELTHTLPI 389  
 QY 333 PSEPHFTLDGNSFYKIIINEERYHICYQ-IDKDCCTFIKGTWTEVI-GIEALTS---DY 388  
 DB 390 PANETFDPRHNGYVDILP-IGYNHLAYFENSNSHYKTLTEGWVNVNGLPADSMENR 448  
 QY 389 LYIISNEYKMGPGGNLYKIQI---SDYTKVTCLSCLNPERCOYYSVPSKEAYYQLR 445  
 DB 449 LYFISTRKSSTE--RHVYIIDLRSPNEIIEVTDISED-----GYVDVSFGSGRRFGILT 500  
 QY 446 CSGQLP---LYTLHS-----SVNDKGLRVLEDNSALDKMLQNVQMPKSKLDFTIL 493  
 DB 501 YKGPVVPQKIVDFHSRAEKCDKGNVLGKSLYHLEKNEVLTKILEDYAVPKSPRELNL 560  
 QY 494 NETKFWYQ-----ILPHFDK--SKYPLLLDVIYAGPCQKADTVFRLNWTYLASTE 545  
 DB 561 GKDBFGKIDLVNSYIELPNDFTLSDHPYFFFAFGPNSSQVVKTSVGFNEVNSQL 620  
 QY 546 NIIIVASPDGSGVQGDKIMHAINRLGTFFVEDOIERAOPSKMGFVNDKRIALWHSY 605  
 DB 621 NAIVVVVDGRTGFGQDFRSLVRDLRGDYEARQISAAISYGLSTFVDPQKISLFGWSY 680  
 QY 606 GGYVTSVVLGSGG-VFKCGIAPVSRWYVYDSVYTERYMGLEPTPEDNLQY-----RN 659  
 DB 681 GGYLTLEKLEKGGHFKYGNVAPVTDWRFYDSVYTERYM--HTPQENFGYVSSVHN 738  
 QY 660 STVMSRAENFKQVEYLLHGTADNVPHPQSAQISKALVDVGVDP-FQAMWYTDDBHGJAS 718  
 DB 739 VTALAAQNR-----LLMHGTGDDNVHPQNSKFLDLIDLNGVENYDVHVPFSDHSIRY 793  
 QY 719 STAHOHYTHMSHFIKQCF 737  
 DB 794 HNANVIVFDKLLDWAQRAF 812

RESULT 10  
 S66261  
 X-Pro dipeptidyl-peptidase (EC 3.4.14.11) - Flavobacterium meningosepticum  
 C:Species: Flavobacterium meningosepticum  
 C:Date: 28-Oct-1996 #sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
 C:Accession: S66261  
 R:Kobashima, T.; Ito, K.; Yoshimoto, T.  
 Arch. Biochem. Biophys. 320, 123-128, 1995  
 A:Title: Cloning, sequencing, and expression of the dipeptidyl peptidase IV gene from *Flavobacterium meningosepticum*  
 A:Reference number: S66261; MUID:95314307; PMID:7793970

A:Accession: S66261  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-711 <KAB>  
 A:Cross-references: UNIPROT:Q47900; UNIPARC:UPI00000B2115; EMBL:D42121; NID:G577283; P  
 C:Superfamily: dipeptidyl-peptidase IV  
 C:Keywords: dipeptidyl-peptidase

Query Match 19.9%; Score 799.5; DB 2; Length 711;  
 Best Local Similarity 29.9%; Pred. No. 2,7e-46;  
 Matches 227; Conservative 131; Mismatches 298; Indels 103; Gaps 27;

QY 10 HHSRKYTYLTYLKYNTYRLK-LYSLRWISDHE-YLYKQNNILVFNABYGNSSVPLENS 67  
 DB 17 HGLSAQETLTKIYSGQYRAKIGSIGSLASLDGENYATIEPTGI---AKSYKTSQKEKN 72  
 QY 68 TPDEFGHSINDYSIPGQPTILLEYVVKWRHSTASYDIYDLNKRQLITEERIPNNTQ 127  
 DB 73 IVD---GSFQGYTFSNDESKILLQKSSQSIYRHSFLGKFEVKDLKSRTWVS---LNNAN 125  
 QY 128 WV---TWSPVGHKLAYVNNNDIYVKIEPNLPSYRITWTGKBDIIVNGITDWTVEEVFSA 184  
 DB 126 WIQPKPSPDGSKVAFIADNNLFQDLNKGITQITTDGKNEIINGLDWYEEB-FGH 184  
 QY 185 YSALWSPNGTFLAYAQFNDTEVPLIEYSFYSDSLQYPKTVRVPYKAGAVNTPKFPV 244  
 DB 185 ADYQWKNKAGDALVFRPDERKVPENIPYIYQN--LYPKLMTYKYPKAGEENSVA--- 238  
 QY 245 VNTDSSLVVTNATSQITAPASMLIGDHYLVDVWAT--QERISLOWLRIONYSVMDIC 302  
 DB 239 --TAYLYQLSSGKSAQLNFGSS---EKYIYIQLFQTNANDEIVVATANRQHN--KVDLL 290  
 QY 303 D-----YDESSGRWNCVLVAROHIEHMTTGWGRPRPSRPHFTLDGNSFYKITSN 351  
 DB 291 KVTYKTAASKLPTETDNW---IETDNLWHE-----FLDONSFL-WASE 331  
 QY 352 BEGYRHICYFOIDKDCCTFIKGTWTEVIGIBALTSYLYISRYKMGPGGNLYKIQLS 411  
 DB 332 RDGHRHLVWYDAAGKLLKQVSKGDWEII-----NYGYNPKTKEYV-IQT 376  
 QY 412 DYTQVTCLSCEL--NPERCOYI-----SVSPSKEAYYQLRCSGPGCLPLYTLHSSVND 462  
 DB 377 EKSINKVWSKLNITGKTQLLSNAEGNSAASFSTFNFTNTSTAKVPTKYILKODANG 436  
 QY 463 KGLRVLDNSALDKMLQNVQMPKSKLDFTIL---NETKFWYQMLPPHDPKSKCYPLL 517  
 DB 437 KDVKELNDDLLNKLKSDNPIAK--EFTIPNAGDQMNW--MKPKNFPDPKCYVVP 492  
 QY 518 LDVYAGPCQKADTVFRLN---WATYLASTENIIVASPDGSGVQGDKIMHAINRLGT 574  
 DB 493 MFQYSGFGSQGVANSWDGGNGIWFMDLAQ-KGYLVVVDGRTGRTGRTGRTKYKVTYKNGLK 551  
 QY 575 FEVEDQIEARQPSKMGFVNDKRIALWHSYGGVYTSVNLGSGSGVFKCGIAPVPSRWE 634  
 DB 552 YEIEDQITAAKWLGNQSYVDKSRIGIFGWSYGGYMASLAMTKGADVFKWGIAPVPTNR 611  
 QY 635 YDVSIVYTERYMGLEPTPEDNLHDYRNSVTMSRAENFKQVEYLLHGTADNVPHPQSAQIS 694  
 DB 612 FYDSIYTERP--LQTPQENKQGYDLNPTTYAKLLKG-KFLIHGTADNVPHPQNSMEFS 668  
 QY 695 KALVDVGVDFQAMWYTDDBHGJASSTAHOHYTHMSHPI 733  
 DB 669 EALIQKQKQDFMAYPDKNHSIIIGNTFRPOLYERKWTNYI 707

RESULT 11  
 T19514  
 hypothetical protein C27C12.7 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T19514  
 R:Thomae, K.  
 submitted to the EMBL Data Library, March 1996



Db 276 LNTGNIBPILPEKSDDDNYELGSLKSLYAHFSPAYNYIFVYENNLFLQQVNSGVAKVY 335  
Qy 161 WTGKEDIYNGITDWMVVEEVPFSAYSALWWSNGTFLAVAOENDETEVPLIEYFYSFYSDESL 220  
Db 336 EDGSD-IFNAKPDWYBEEVLASQAIAWAPDDSKAVFARFNDISVDIRLNRNTYNNWE 394  
Qy 221 QYPTKTRVYPKAGAVNPVTFVFNVTDSLSVTNATSIQITAPASMLIGHYLCADVTA 280  
Db 395 AYLSDTKIKYKPGFQNFQDLFLVN-----LQNGIYISINTGQK---DSILYNGKI 445  
Qy 281 TQERISLQWLRIQNYVMDICDYDESSGRWNLVARQHIEMSTTGWRFR-----PSE 335  
Db 446 SPDTFRFRITDR--NSKILDVKVYDIPSSQ--MLTVRNTNSLNFNGIEKTKDILSIPP 501  
Qy 336 PHFTLDGNSFYKIIISNEEGYRHICYF-OLDKDCCTFIFKGTWEV-----IGIEALTSYVL 389  
Db 502 PELKEMDGYDIHADSRGFSHLFYPTVFAKEPIQLTKGWEVTGNGIVGEYET-DTI 560  
Qy 390 YYISNEYKMPGGRNLYKIQLSDYTKVCLSCENLP-BRCQYYSVSFSKEAKYQLRCSG 448  
Db 561 FFTANEIGVM--SQHLYSISLTDSTTQNTFQSLQNPDKYDYPDELSSARYAISKLG 618  
Qy 449 PGLPLYTH-----SSVNDKGLRVLEDSALDKMLQNVOMPSKKLDIFILNE-TKFM 499  
Db 619 PDTEPIKAVPLTRVLNVAEIHDDSLQLTKDEKFEKIKNYDLPTITSYKTMVLDGVEIN 678  
Qy 500 YOMILPPHFDSKYPILALDVYVAGPCOKADTVFRLNWTATVLASTENIIVASFQDGRSGY 559  
Db 679 YIEIKPANLNPKKYPILVNIYGGSGQTFTTKSLAFAEQAVVSLDVLVILQIEPRGTGG 738  
Qy 560 QGDKIMHAINRRLGTFEVEDQTEAARQFSKMG--FVDNKRITAIWGSYGGVYVTSMLVLSG 617  
Db 739 KGMSPRSWAREKLGWEPRDITEVTKFIQNSQHIDSKIAIWGSYGGTSLKTVELD 798  
Qy 618 SG-VFKGCIAPVSRWEYDSVYTERVGLPTPEDNLDHNRNSTVMSRAENFKQVEYL- 675  
Db 799 NGDTFKYAMAVAPNTWLYSDVYTERYMNQPS-ENHEGYFEVSTI-----QNFKSFESLK 853  
Qy 676 --LIHGTDNNVHFQOQAQISKALVDVGV--DFQAMWYTDDEHGIASSTAHOIYTHMSH 731  
Db 854 RLPVHGTFDNNVHIONTRFLVDQLNLGLTNMDHIIFFDSHDSIRYHNAQRIVFQKUY 913  
Qy 732 FIKQCF 738  
Db 914 WLRDAFA 920

RESULT 13  
T25174  
hypothetical protein T23F1.7b - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T25174  
R;Wilkinson, J.  
submitted to the EMBL Data Library, October 1996  
A;Reference number: 219990  
A;Accession: T25174  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-799 <WIL>  
A;Cross-references: UNIPROT:O18119; UNIPARC:UPI000061190; EMBL:281129; PIDN:CAB03412.1;  
A;Introns: 13/3; 52/1; 111/3; 197/3; 366/3; 524/1; 557/1; 648/2; 703/1; 743/1  
C;Genetics:  
A;Gene: CRSP:T23F1.7b  
A;Map position: 5  
A;Superfamily: dipeptidyl-peptidase IV  
Query Match 18.0%; Score 725; DB 2; Length 799;  
Best Local Similarity 27.4%; Pred. No. 3.7e-41;  
Matches 215; Conservative 119; Mismatches 282; Indels 168; Gaps 31;

Qy 14 RKTVLTLDYLNK-----TYRLKLYSLRWISDHEYLKQENNLVFNFAEYGNSSVFLPE 65  
Db 120 RKYFAMQDHAHPGNWNPQNETFHLKI-----VNNNERL-----NPLLPEVE----- 161  
Qy 66 NSTDFEFGHSINDYSISPDGQFILLVYVQWRHSYASDYIDLNRQLITBERIPNN 125  
Db 162 -----ELFRELSDSRIT-----YDIGLRK-----BESV----- 184  
Qy 126 TQWVTSVPVGHKLAVWNNDIYVKIEPNLPSYRIWTGCKEDIYNGITDWMVVEEVPFSAYS 185  
Db 195 IQAFKXNGKDFVVFESNKLIYQSSPEEGLTRVNSGGEHTV-DGLDFWIEEIEIFGRK 243  
Qy 186 SALMWSNGTFLAVAQFNDTEVPLIEYFYS-----DESLEYKPTKTVVPYKAGAVNP 238  
Db 244 DAMWSTKGQDLAVASVDNHLTKNLSVLTYHRLBEPIDTNFHYPKTF-----AKVLP 296  
Qy 239 TVKFPVNTDSLSVTNATSIQITAPASMLIGHYLCADVTA-----ATQERISLQWLRIQ 294  
Db 297 TY-----TLSIWNKTKESQRLDQVLKDSLSYHLLAVKWLBEINGTQLVSV-WTNRVQ 349  
Qy 295 NYSVMDICDYDESSGRWNLVARQHIEMSTTGWRFRPSEPHFTLDGNSFYKIIISNE 352  
Db 350 NEVALTICD-----MDTAICRLEFFEYKYASKRWVTH-----DDFHISITFE 390  
Qy 353 EGYRHCYFQI--DKKDC-----FITGTWEVIGIEALTSYLYI 392  
Db 391 D-----TLFLLPHDKRDNAFQOQVSLRSLSHGQLRTPKFLNLGEYDVTSGINGKETRTIF 446  
Qy 393 SNEYKMPGGRNLYKIQLSDYTK--VTCLSCENLPBRCQYYSVSFSKEAKYQLRCSPG 450  
Db 447 FHAAAPKPSHRSLFSYSLADESRNSAYCISCSI--KNCTWAAQAMDDQMKTAIVSCKGA 504  
Qy 451 LPLYLTHSSV-----NKGRLVLEDSALDKMLQNVOMPSKKLDIFILNETKF-----W 499  
Db 505 AP-----HTAIVNTRFMSDDKKEHANLLYDKTYQN-RVEEAGLPVVIKETIKISDDFDAL 559  
Qy 500 YOMILPPHF-----DKSKYPILALDVYVAGPCOKADTVFRLNWTATVLASTENIIVASFQDGRG 556  
Db 560 IKLSIPKDIYNRDKHQAIPLIVHVYGGPNDQTKTEATQIGIEEVVASASQAAILRIDGRG 619  
Qy 557 SGYQGDKIMHAINRRLGTFEVEDQTEAARQFSKMG--FVDNKRITAIWGSYGGVYVTSMLVLSG 613  
Db 620 SGGRGWYRSAYIYQQLGTVEVEDQIKAKVVLRYRLHLLDARRVAVFGWSYGGPMTLSMV 679  
Qy 614 LGSGGVFKGCIAPVSRWEYDSVYTERVGLPTPEDNLDHNRNSTVMSRAENFKQVE 673  
Db 680 NEAPEQFCKAVSAPVNFAYDTYTERVGLPTPEDNLDHNRNSTVMSRAENFKQVE 733  
Qy 674 YLLIHGTADDNNVHFQOQAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHF 733  
Db 734 LLLMHGLDDNNVHFQNSAILIDELQNRGVDFDLVYVNPQNSLSRSTS--HVVVGWTHFL 791  
Qy 734 KQCF 737  
Db 792 RQCF 795

RESULT 14  
T25173  
hypothetical protein T23F1.7a - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T25173  
R;Wilkinson, J.  
submitted to the EMBL Data Library, October 1996  
A;Reference number: 219990  
A;Accession: T25173  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-779 <WIL>  
A;Cross-references: UNIPROT:O18119; UNIPARC:UPI000002A220; EMBL:281129; PIDN:CAB03411.1;  
A;Experimental source: clone T23F1  
C;Genetics:

[illegible]

RESULT 15  
A87516  
dipeptidyl peptidase IV [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
C:Accession: A87516  
C:Jernigan, W.C.; Feldblivum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B. J. Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolodziej, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 96, 4136-4141, 1999  
A>Title: Complete Genome Sequence of *Caulobacter crescentus*.  
A|Reference number: A87249; PMID:21173698; PMID:11259647  
A|Accession: A87516  
A|Status: preliminary  
A|Molecule type: DNA  
A|Residues: 1-738 <STO>  
A|Cross-references: UNIPROT:Q9A5E0; UNIPARC:UPI00000C7616; GB:AE005673; NID:G13423647; P  
C|Genetics:  
A|Gene: CC2154

Query Match	15.6%	Score 628;	DB 2;	Length 738;
Best Local Similarity	26.8%;	Pred. No. 1.2e-34;		
Matches 189;	Conservative 115;	Mismatches 269;	Indels 132;	Gaps 23;
Qy	76	INDYSISPOQFILLVY	--YVQWRHSY	TA SYD YD LNK QRL I I T E E R I P N N T Q W T W S 132
Db	111	I V E Y S D R Q R F I L V P L D G D L Y	-----DAVADGK	I T R L I E --T P G D E V D A K V S 157
Qy	133	P V G H K L A Y W N N D I Y V K I E N P L S Y R I T W G K E D I I Y N G I T D W V Y E E E V S F A S Y A L W M S P 192		
Db	158	P K G Y V S Y R D Q N L Y I K P V A G G A E T A L T T D G K A L S F - G V A E F I V Q E L - D R E T G Y W M S P 215		
Qy	193	N G T F L A Q A N D T E V P L I E V S F S D S L Q P K T V R P Y P K A G A V N T P K F F V N T D S L S S 252		
Db	216	D E S R I V Y T R D E S G D I V P - - - - - R A D I P G G G A T V V N Q R Y P R A G R P N A V D L F - - - - - 263		
Qy	253	V T N A T S I Q I T A P A S M L I G D H Y L C D V T W - A T Q E R I S L Q W L R R I Q N Y S V M D I C D Y D E S S G R W 311		
Db	264	V R L A S G K V T A L D L G A N K D I Y A R V A W S A D C K T V Y V Q R L S R D Q - K T L D L L A F D A T G A G 321		
Qy	312	N C L V A R G H I E M S T T G W G R P S P E H F T L D G N S F Y K I I - - - - - S N E G Y R H I C Y P Q I D 364		
Db	322	K T I L T - - - - - D T D P H F I E V S N D F R P L T D G T F L W G S E K D Q N H L Y R Y A A D 365		
Qy	365	K K O C T F I T K T W E V I G I E A L T S D Y L Y I S N E Y K G M P G G R N L Y K I Q L S D Y T K V T L C S E L N 424		
Db	366	G K L I A T I T K D W P V I G L E G - - - - - V D E A R K V A I F S A S I D 399		
Qy	425	- P E R C O Y V S Y S F S K E A K Y I O L R C S G - - - - - P G L P Y L T L H S S V N D K 463		
Db	400	T P I E R R L Y E S Y A K P G K P K A L T S A G G W A A K A V A D N G A F A G T Y S D P K T P S Q T A L Y S A D G K 459		
Qy	464	G L R V L E D N S A L D - - - - - K M L Q N Y Q M P S - - - - - K K L D F I I L N E T K F W Y Q M I L P P H F D K S K Y 514		
Db	460	R V R W I E N K L A E G H P Y P Y A A N L P Q E F G S L K A A D - - - - - G E T - L H Y E I L K P I G F D P A K Y 514		
Qy	515	P L L D V Y A G P C S Q A D T V F R L N W - - - - - A T Y L A S T E N I I V A S P D R G S G V G Q G K I M A I N 569		
Db	515	P A I V S Y V G G P H A Q R V - - - - - M K N W H S P E R T Y L E A - - - - - G Y V I F K L D N R G S G R S A K F M R A L D 568		
Qy	570	R R L G T F E V E D I E A A R Q F S K W G F V D N K R I A I W G S Y G G V Y T S V M L G S G S G V F K G I A V A P 629		
Db	569	R K L G T F E V E D L L G A K F L A S Q P Y V D A D K L G V M G S Y G G F M A L M L L T A E N T P F F K A G A C A P 628		
Qy	630	V S R W E Y D S V V T E R Y M G L P T P E D N L H Y R N S T Y M S R A E N F K Q V E Y L L I H G T A D D N V H F Q 689		
Db	629	P T E W S L Y D T A Y T E R Y M G - - - - - K P D E N K A G A Y A S D I N N R I D K L A P G S L L L L H G W A D D N V I F E N 686		
Qy	690	S A Q I S K A L V D V G D V F Q A W M Y T D E D H G I A S T A H - - - - - Q H I Y T M H S H F 732		
Db	687	S T R L M A L O R K A I L F E N A M Y P G E R H S A P G S K T K G L S V L K T H L D P F F 731		

Search completed: February 15, 2006, 13:15:27  
Job time : 48 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2006, 13:26:39 ; Search time 18 Seconds  
(without alignments)  
584.259 Million cell updates/sec

Title: US-10-659-055-3

Perfect score: 4020

Sequence: 1 ADPCGSHHHHSRKYTLT.....AQHIYTHMSHFIRKQCSLP 740

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 107799 seqs, 14211699 residues

Total number of hits satisfying chosen parameters: 107799

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:

- 1: /cgn2\_6/ptodata/2/pubaa/US08\_NEW\_PUB pep.\*
- 2: /cgn2\_6/ptodata/2/pubaa/US06\_NEW\_PUB pep.\*
- 3: /cgn2\_6/ptodata/2/pubaa/US07\_NEW\_PUB pep.\*
- 4: /cgn2\_6/ptodata/2/pubaa/PCT\_NEW\_PUB pep.\*
- 5: /cgn2\_6/ptodata/2/pubaa/US09\_NEW\_PUB pep.\*
- 6: /cgn2\_6/ptodata/2/pubaa/US10\_NEW\_PUB pep.\*
- 7: /cgn2\_6/ptodata/2/pubaa/US11\_NEW\_PUB pep.\*
- 8: /cgn2\_6/ptodata/2/pubaa/US60\_NEW\_PUB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3939	98.0	766	6	US-10-522-789-2
2	3886	96.7	762	7	US-11-116-939-13
3	2168	53.9	760	7	US-11-186-284-55
4	550	13.7	102	7	US-11-176-951-7
5	528.5	13.1	882	7	US-11-151-601-20
6	410.5	10.2	102	7	US-11-176-951-8
7	291	7.2	54	7	US-11-179-977-1
8	272	6.8	657	7	US-11-179-977-1
9	167	4.2	115	7	US-11-176-951-9
10	167	4.2	115	7	US-11-176-951-12
11	145	3.6	27	7	US-11-176-951-6
12	145	3.6	27	7	US-11-176-951-13
13	145	3.6	109	7	US-11-176-951-10
14	130	3.2	99	7	US-11-176-951-11
15	124	3.1	668	6	US-10-454-437-118
16	116.5	2.9	1593	6	US-10-453-372-1092
17	114	2.8	960	7	US-11-169-041-177
18	114	2.8	2516	6	US-10-647-956A-2
19	112.5	2.8	1436	6	US-10-453-372-1094
20	110	2.7	19	7	US-11-116-939-27
21	110	2.7	72	7	US-11-151-601-22
22	110	2.7	453	7	US-11-024-959-355
23	109.5	2.7	877	7	US-11-077-550-157
24	107.5	2.7	3194	7	US-11-052-554A-90
25	107	2.7	866	7	US-11-077-550-32

26	107	2.7	871	7	US-11-077-550-155	Sequence 155, App
27	107	2.7	873	7	US-11-077-550-163	Sequence 163, App
28	107	2.7	873	7	US-11-077-550-165	Sequence 165, App
29	107	2.7	876	7	US-11-077-550-66	Sequence 66, Appl
30	107	2.7	878	7	US-11-077-550-62	Sequence 62, Appl
31	107	2.7	879	7	US-11-077-550-30	Sequence 30, Appl
32	107	2.7	908	7	US-11-077-550-64	Sequence 64, Appl
33	107	2.7	914	7	US-11-077-550-60	Sequence 60, Appl
34	107	2.7	949	7	US-11-077-550-68	Sequence 68, Appl
35	107	2.7	1130	7	US-11-077-550-44	Sequence 44, Appl
36	107	2.7	1132	7	US-11-077-550-46	Sequence 46, Appl
37	107	2.7	4495	6	US-10-453-372-1002	Sequence 1002, Ap
38	106	2.6	878	7	US-11-077-550-12	Sequence 12, Appl
39	106	2.6	907	7	US-11-077-550-16	Sequence 16, Appl
40	106	2.6	953	7	US-11-077-550-14	Sequence 14, Appl
41	106	2.6	1013	7	US-11-077-550-18	Sequence 18, Appl
42	105.5	2.6	305	7	US-11-098-686-10980	Sequence 10980, A
43	105.5	2.6	658	6	US-10-873-528-17	Sequence 17, Appl
44	105.5	2.6	677	6	US-10-873-528-155	Sequence 155, App
45	105.5	2.6	871	7	US-11-077-550-2	Sequence 2, Appl

#### ALIGNMENTS

##### RESULT 1

US-10-522-789-2  
; Sequence 2, Application US/10522789  
; Publication No. US20050260732A1  
; GENERAL INFORMATION:  
; APPLICANT: TANABE SEIYAKU CO., LTD.  
; TITLE OF INVENTION: Three-dimensional structure of dipeptidyl peptidase IV  
; FILE REFERENCE: 03-039-PCT  
; CURRENT APPLICATION NUMBER: US/10/522,789  
; CURRENT FILING DATE: 2005-01-28  
; PRIOR APPLICATION NUMBER: US 60/398,761  
; PRIOR FILING DATE: 2002-07-29  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 766  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-522-789-2

Query Match 98.0%; Score 3939; DB 6; Length 766;  
Best Local Similarity 100.0%; Pred. No. 2.2e-298; Mismatches 0; Indels 0; Gaps 0;  
Matches 728; Conservative 0;

Qy	13	SRKTYTLTDYLNKNTYRLKLSLRWISDHEVLYKQENILVFNABYGNSSVPLENSTPDEF	72
Db	39	SRKTYTLTDYLNKNTYRLKLSLRWISDHEVLYKQENILVFNABYGNSSVPLENSTPDEF	98
Qy	73	GHSINDYSISPDGQFILLENNYVQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVWS	132
Db	99	GHSINDYSISPDGQFILLENNYVQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVWS	158
Qy	133	PVGHKLAVVWNNDIYVVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVPSAYSALWSP	192
Db	159	PVGHKLAVVWNNDIYVVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVPSAYSALWSP	218
Qy	193	NGTFPLAQAQNDIEVPLIEYSFYSDLSQPKTVRPVYPKAGAVNPTVKPFVWNTDSLSS	252
Db	219	NGTFPLAQAQNDIEVPLIEYSFYSDLSQPKTVRPVYPKAGAVNPTVKPFVWNTDSLSS	278
Qy	253	VTNATSIQTAPASMLIGDHYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWN	312
Db	279	VTNATSIQTAPASMLIGDHYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWN	338
Qy	313	CLVARQHIEMSTGTCWGRFRPSPHFTLDGNSFYKIIISNBEGRHICVFQIDKKDCTFIT	372
Db	339	CLVARQHIEMSTGTCWGRFRPSPHFTLDGNSFYKIIISNBEGRHICVFQIDKKDCTFIT	398

373	Qy	KGTWEVIGIEALTS	DYLYIYN	SNEYKMPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS	432
399	Db	KGTWEVIGIEALTS	DYLYIYN	SNEYKMPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS	458
433	Qy	VSFSEAKYYQLRCSG	RGLPIYLTHSSVNDKGLRVLEDN	SALDKMLQNVMSPSKLDPLII	492
459	Db	VSFSEAKYYQLRCSG	RGLPIYLTHSSVNDKGLRVLEDN	SALDKMLQNVMSPSKLDPLII	518
493	Qy	LNETKFWMYOILP	PHFDKS	KKYPLLDDVYAGPCSQKADTVFRLNWATYLASTENIIVASF	552
519	Db	LNETKFWMYOILP	PHFDKS	KKYPLLDDVYAGPCSQKADTVFRLNWATYLASTENIIVASF	578
553	Qy	DGRSGYGQD	KIMHAINRR	LFTFEVBQIEAARQFSXMGFVDNKRIALMGWSVGGVYTSM	612
579	Db	DGRSGYGQD	KIMHAINRR	LFTFEVBQIEAARQFSXMGFVDNKRIALMGWSVGGVYTSM	638
613	Qy	VLGSGSGVPKC	GIAVAPSRWEYDSVYTTERYMGLPTPEDNL	DHYRNSTVMSRAENPKOV	672
639	Db	VLGSGSGVPKC	GIAVAPSRWEYDSVYTTERYMGLPTPEDNL	DHYRNSTVMSRAENPKOV	698
673	Qy	EYLLHTHG	TADDNVHFQQSAQISKALVDVGVDFQAMWTTDEDHGIASSTAHOHIYTHMSHF	732	
699	Db	EYLLHTHG	TADDNVHFQQSAQISKALVDVGVDFQAMWTTDEDHGIASSTAHOHIYTHMSHF	758	
733	Qy	IKOCPSLP	740		
759	Db	IKOCPSLP	766		

## RESULT 2

```

US-11-116-939-13
; Sequence 13, Application US/11116939
; Publication No. US20050265995A1
; GENERAL INFORMATION:
; APPLICANT: Stephen Tomlinson
; APPLICANT: Richard J. Quigg
; TITLE OF INVENTION: TISSUE TARGETED COMPLEMENT MODULATORS
; FILE REFERENCE: 19113.011502
; CURRENT APPLICATION NUMBER: US/11/116,939
; CURRENT FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: 60/565,907
; PRIOR FILING DATE: 2004-04-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 762
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; note=synthetic
; OTHER INFORMATION: construct
US-11-116-939-13

```

Query Match	96.7%;	Score 3886;	DB 7;	Length 762;
Best Local Similarity	99.0%;	Pred. No. 2.8e-294;		
Matches 721; Conservative	0;	Mismatches 3;	Indels 4;	Gaps 1
Qy	13	SRKTYTLDTYLNKTYRLKLYSLRISDSDEHYLYKQNNILVFNAEYGNSSVPLENSFDFSF	72	
Db	39	SRKTYTLDTYLNK---YLKCLRWSDSDEHYLYKQNNILVFNAEYGNSSVPLENSFDFSF	94	
Qy	73	GHSINDYSIPDQFILLLEYNVYVKWRHSYTYSDIYDLNKRQLITSEIRIPNNQTQWTVWS	132	
Db	95	GHSINDYSIPDQFILLLEYNVYVKWRHSYTYSDIYDLNKRQLITSEIRIPNNQTQWTVWS	154	
Qy	133	PVGHKLAUVWNNDIYVKLEPNLPSYRITWTGKEDIITNGITDMWYEEVEFSAYSLAWNPS	192	
Db	155	PVGHKLAUVWNNDIYVKLEPNLPSYRITWTGKEDIITNGITDMWYEEVEFSAYSLAWNPS	214	
Qy	193	NGTFLAYAOFNDTEVPLIEYSPYSDESIYQPKTVRVPYPKAGAVNPVTKPFVVNTDSLSS	252	
Db	215	NGTFLAYAOFNDTEVPLIEYSPYSDESIYQPKTVRVPYPKAGAVNPVTKPFVVNTDSLSS	274	

Qy	253	V	T	N	A	T	S	I	Q	I	T	A	P	A	S	M	L	I	G	D	H	L	C	D	V	T	W	A	T	O	E	R	I	S	L	O	W	L	R	I	Q	N	Y	S	V	M	D	I	C	D	Y	E	S	S	G	R	W	N	312	
Db	275	V	T	N	A	T	S	I	Q	I	T	A	P	A	S	M	L	I	G	D	H	L	C	D	V	T	W	A	T	O	E	R	I	S	L	O	W	L	R	I	Q	N	Y	S	V	M	D	I	C	D	Y	E	S	S	G	R	W	N	334	
Qy	313	C	L	V	A	R	O	H	I	E	M	S	T	G	W	G	R	F	P	S	P	H	F	L	D	G	N	S	F	V	K	I	I	S	N	E	E	G	R	H	I	C	F	O	I	D	K	D	C	T	F	I	372							
Db	335	C	L	V	A	R	O	H	I	E	M	S	T	G	W	G	R	F	P	S	P	H	F	L	D	G	N	S	F	V	K	I	I	S	N	E	E	G	R	H	I	C	F	O	I	D	K	D	C	T	F	I	394							
Qy	373	K	G	T	M	E	V	I	G	I	E	A	L	T	S	D	Y	L	I	Y	I	S	N	E	Y	K	G	M	P	G	G	R	N	L	Y	K	I	Q	L	S	D	Y	T	K	V	T	C	L	S	C	E	L	N	P	E	R	C	Q	Y	432
Db	395	K	G	T	M	E	V	I	G	I	E	A	L	T	S	D	Y	L	I	Y	I	S	N	E	Y	K	G	M	P	G	G	R	N	L	Y	K	I	Q	L	S	D	Y	T	K	V	T	C	L	S	C	E	L	N	P	E	R	C	Q	Y	454
Qy	433	V	S	F	S	K	E	A	K	Y	Q	L	R	C	S	G	P	L	P	L	T	H	S	V	N	D	K	G	L	R	V	L	E	D	N	S	A	L	D	K	M	L	O	N	V	M	P	S	K	L	D	P	I	492						
Db	455	V	S	F	S	K	E	A	K	Y	Q	L	R	C	S	G	P	L	P	L	T	H	S	V	N	D	K	G	L	R	V	L	E	D	N	S	A	L	D	K	M	L	O	N	V	M	P	S	K	L	D	P	I	514						
Qy	493	L	N	E	T	K	F	W	Y	Q	M	L	P	P	H	F	D	K	S	K	Y	P	L	L	D	V	A	G	P	C	S	Q	K	A	D	T	V	F	R	L	N	W	A	T	Y	L	A	S	T	E	N	I	V	A	S	F	552			
Db	515	L	N	E	T	K	F	W	Y	Q	M	L	P	P	H	F	D	K	S	K	Y	P	L	L	D	V	A	G	P	C	S	Q	K	A	D	T	V	F	R	L	N	W	A	T	Y	L	A	S	T	E	N	I	V	A	S	F	574			
Qy	553	D	G	R	G	S	G	Y	Q	G	D	K	I	M	H	A	I	N	R	L	G	T	E	V	E	D	Q	I	E	A	A	R	Q	F	S	K	M	G	F	V	D	N	K	R	I	A	I	G	W	S	Y	G	G	Y	T	S	M	612		
Db	575	D	G	R	G	S	G	Y	Q	G	D	K	I	M	H	A	I	N	R	L	G	T	E	V	E	D	Q	I	E	A	A	R	Q	F	S	K	M	G	F	V	D	N	K	R	I	A	I	G	W	S	Y	G	G	Y	T	S	M	634		
Qy	613	V	L	A	S	G	S	G	V	F	K	G	I	A	V	P	S	R	W	E	Y	D	S	V	Y	T	R	Y	N	G	L	P	T	P	E	D	N	L	D	H	R	N	S	T	V	M	S	R	A	E	N	F	K	Q	V	672				
Db	635	V	L	A	S	G	S	G	V	F	K	G	I	A	V	P</																																												



```

; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 7
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-7

Query Match      13.7%; Score 550; DB 7; Length 102;
Best Local Similarity 100.0%; Pred. No. 3.9e-36;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PGGSHHHHKKYTTLDYLNKTVRLKLYSLRWISDHELYKQ--ENNILVFNAEYGNSS 61
Db 28 PSRVHSENTRWALTLDKILNGTSYKTFPPNWLISGQEYLLHQSDANNILVLTNIITGQSY 87

QY 62 VFLENSTDFBGHSIN--DYSISPDGQFILLEYNVYKWRHSYTYASDIYDLNKRQLITE 119
Db 88 TILSNRWT---KSVNASNYGLSPDRQFVYLESYSLKWRYSYTTATYYIYDLSNGEFVRG 143

QY 120 ERIPNNTQWVWSPGHKLAYWNDIYVKLEPNLPSVRIITWTKGEDIIYNGITDWYVEE 179
Db 144 NELPAPIQYLCWSPGVSKLAVYQNNIYLKQRPDPPFPQITFNGRENKIFNGIPDWYVEE 203

QY 180 EYFSAYSALWSPNGTFLAYAQNDTEVPLIEYSFYSDSLQYPKTVRPVPPKACAVNPT 239
Db 204 EMLPTKYALWSPNGKFLAYAEFNDKDIPTVAYSYGDE--QYPTNIPIPPKAGAKNPV 261

QY 240 VKFFVNTDSLSVNTAISIQTAPASMLIGDHYLCDVTWATQERISLQWLRRIQNYSVM 299
Db 262 VRIFIIDTYPAYVGPQ---EVPVPMIASSDYFSLTWTWDERVCLQWLKRVQNVSVL 318

QY 300 DICDYDESSGNCLVARGHEIMSTGWGVRPSEPHEFTLDGNSFYKLIINEBGRHIC 359
Db 319 SICDFREDWQWDCPKTQEHIEESRTGWAGGFVSRPVPFSYDAISYKIFSDKQGYKIH 378

QY 360 YFQDIDKDCITTKGTWVIGIEALTSYLVYISNEYKMGPRGNRLYKIQLSDY--TKYTC 418
Db 379 YIKDTVENAIQITSGKWEAINIFRTQDSLFISSNEFEYPPRRNIYRISIGSVPPSKKC 438

QY 419 LSCELNPERCOYYSVSFEKARYQLRCSGGLPLTYLLHSSVNDKGLRVLEDNSALDRL 478
Db 439 VTCHLRKERCQYTTASPSDYAKYALVCYGGIPISTILDGRTDQEIKILEENKELENAL 498

QY 479 QNVQMPSKGLPIILNETKFWYQMLPPHPDKSKYPIILLVDYAGPCQKADTVFRLNWA 538
Db 499 KNIQLPKEEIKLEDEITLWYKMLPPQDRSKYPIILLIQQYGGPCQSVRSFVAVNVI 558

QY 539 TYLASTENIIVASPDGRSGVGQDKIMHAINRRLGTPEVEDQIEAARQFSKMGVDNCKRI 598
Db 559 SYLASKEGNVIALVDGRTACQGDLLYAVTRKLVGIVEVDQITAVRKFIEMGFIDEKRI 618

QY 599 AIWGSYGYVTSVVLGSGGVFKGIAVAPVSRWEYDYSVYTERYMGLETPEDNLDHYR 658
Db 619 AIWGSYGYVSSIALASGTGLFKGIAVAPVSSWEYASVYTERFMSLPTKDDNLEHYK 678

QY 659 NSTVMSRAENFKQVEYLLIHGTADNDNVHFQOQSAISKALVDVGVDFQAMWTTDEDHGAS 718
Db 679 NSTVMARAEBYFRNVDYLLIHGTADNDNVHFQOQSAIKALVNAQVDVQAMWYSDQNHGL-S 737

QY 719 STAQHIYTHMSHFIKQCFSL 739
Db 738 GLSTNHLIYTHMTHFLLKQCFSL 758

RESULT 4
US-11-176-951-7
; Sequence 7, Application US/11176951
; Publication No. US20060024313A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, YUAN-SHOU
; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
; FILE REFERENCE: 08842.0019
; CURRENT APPLICATION NUMBER: US/11/176.951
; CURRENT FILING DATE: 2005-07-06
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952
; PRIOR FILING DATE: 2004-07-06
; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 7
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-7

Query Match      13.1%; Score 528.5; DB 7; Length 882;
Best Local Similarity 24.4%; Pred. No. 3.4e-33;
Matches 210; Conservative 121; Mismatches 339; Indels 191; Gaps 40;

QY 2 DPGSGHHH-----HHSRKTYTLDYLNKTVRLKLYSLRWISDHELYKQENNILVF 53
Db 74 DPGPHSDRIYVYLAAMSGENRENTLTFSEIPKTNRAAVLMSW-----KPLLDLFOA 125
```

```

; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 7
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-7

Query Match      13.7%; Score 550; DB 7; Length 102;
Best Local Similarity 100.0%; Pred. No. 3.9e-36;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 639 VYTERYMGLETPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADNDNVHFQOQSAISKALV 698
Db 1 VYTERYMGLETPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADNDNVHFQOQSAISKALV 60

QY 699 DVGVDVQAMWVTTDEDHGASSTAQHIYTHMSHFIKQCFSLP 740
Db 61 DVGVDVQAMWVTTDEDHGASSTAQHIYTHMSHFIKQCFSLP 102

RESULT 5
US-11-151-601-20
; Sequence 20, Application US/11151601
; Publication No. US20060003413A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williams, Mark
; APPLICANT: Welch, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE.
; FILE REFERENCE: MF100-054FPCIPIONIDIVIM
; CURRENT APPLICATION NUMBER: US/11/151.601
; CURRENT FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: US 10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-151-601-20

Query Match      13.1%; Score 528.5; DB 7; Length 882;
Best Local Similarity 24.4%; Pred. No. 3.4e-33;
Matches 210; Conservative 121; Mismatches 339; Indels 191; Gaps 40;

QY 2 DPGSGHHH-----HHSRKTYTLDYLNKTVRLKLYSLRWISDHELYKQENNILVF 53
Db 74 DPGPHSDRIYVYLAAMSGENRENTLTFSEIPKTNRAAVLMSW-----KPLLDLFOA 125
```

```
QY 54 NAEYG-----NSSVFLNENSTFDFGHSINDYSISPDQOFILLEVNVYKQHRHSYTSYDIY 109
Db 126 TLDYGMYSREELLRRKRGITGVGIASYDYH-QSGTFL-----FQAGSGIY 171
QY 110 DLNK-----RLQITEIRIPN-----NTQVVTWSPVGHKLAYVWNNDIYV 148
Db 172 HVKGGPQGFQTOPLRPNLVEISCPNRMDDPKLCPADPMI-----AFIHSNDIWI 222
QY 149 KIEPNLPSVRIWTG-----KEDIYNGITWVYEEVFSAYSALWSPNGTF-----196
Db 223 SNIVTREERLLTVVHNLAMBEDARSAGVATFVLQEE-FDRYSGYWMCPKAETTPSGGK 281
QY 197 ---LAAQNDTEVEPLIEVSFSDLESLOYPKTVRYPKAGAVNPTVKEFV---VNTDLS 251
Db 282 IIRILYEENDESEVEIHH---VTSPLMETRADSPRYPKTGANTPKVTFKMESEIMIDAG 338
QY 252 SVTNATSIQITAPASMLI-GDHYLCDVTWA-----TOERISLOWL-----290
Db 339 RIIDVIDKELIQPELIFEGVEYIARAGWTPSGKYAWSILLDORSQRLQIVLISPFLFTP 398
QY 291 -----RRIONY---SVMDICDVDESSGRM-NC-----LVARQHIEMSTTGWGRFRP 333
Db 399 VEDDVMERQLIESVPDSVTPLIYEETTDIWINIHDFHPVFPQSHEEIEIFASECKT 458
QY 334 SEPHTLDGNSFKIIS---NEEGYRHI-----CYFQIDKKDCTFITKGTWEVIG---380
Db 459 GFRH-----LYKITSILKESKXSGGLPAPSDFKCPKIEEIAITSGEWELGRHGS 511
QY 381 -JEALTSYLYISNEYKMPGGRNLYKI-----QLSD---YTKVTCLSCELNPERC 428
Db 512 NQOVDEVRLLVPEGT-KDSPLEHLLYVSYVNPGEVTELTDRGYSHSCIS-----QHC 565
QY 429 QYYSVSFSKAKYQLRCSGGLPLPYLTHSSVNDKGLRVLED-NSALDKM---LQNVQWPS 485
Db 566 DPFISKYSQKNPH---C---VSLYKLSPPDDPTCKTKFEMATILDSAGPLPDVTPPE 618
QY 486 KKLDPILLNETXF-WQMLPLPH-PDKSKKYPILLDVVAGPCSQ-----KADTVFRLNW 537
Db 619 ---IFSFESTTGFTLYGMLYKPHDQPGKKYPTVLFIYGGPQVLVNNRFGVKYFRLN- 674
QY 538 ATYLASTENIIVASDFGRSGYQGDKIMHAINRRLGTPEVEQIEAARQF-SKMGFVDNK 596
Db 675 ---TLASLGIVVVV-IDNRSGCHRLGKFEGAFKYKMGQIEIDQVEGLQVLAIRYDFIDL 731
QY 597 RIALMGWSYGGVYTSVNLGSGGVFKCGIAPVPSRWBYSDSYTYTYRYMGLTPEDNLDH 656
Db 732 RVGIHGSYGGVLSLMALMQRSDFRVALAGAPVTLWIFDYTGTYRYMG--HPDQNEQG 789
QY 657 YRNSTVMSAENF--KQVEYLLIHGTADDNVHVFQOQAISKALVDVGVDFQAMWYTDDEH 714
Db 790 YVLGSVMAQAEKPPSPENRLLHGLFDENVHFAHTSILLFLVRAGKPYDILQIYPOERH 849
QY 715 GIASSTAHOHIYTHMSHFIKQ 735
Db 850 SIRVPESGEHYELHLLHYLQE 870
```

## RESULT 6

```
US-11-176-951-8
; Sequence 8, Application US/11176951
; Publication No. US20060024313A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, XIN
; APPLICANT: CHEN, YUAN-SHOU
; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
; TITLE OF INVENTION: PROLYL DIPEPTIDASES
; FILE REFERENCE: 08842.0019
; CURRENT APPLICATION NUMBER: US/11/176,951
; CURRENT FILING DATE: 2005-07-06
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952
```

```
; PRIOR FILING DATE: 2004-07-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 8
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-8
```

```
Query Match 10.2%; Score 410.5; DB 7; Length 102;
Best Local Similarity 72.3%; Pred. No. 2.7e-25;
Matches 73; Conservative 17; Mismatches 10; Indels 1; Gaps 1;
```

```
QY 639 VYTERYMGLPPTPEDNLDHVRNSTVMSRAENFKQVEYLLIHGTADDNVHVFQOQAISKALV 698
Db 1 VYTERYMGLPPTKDDNLEHYKNSVWRAEYFRNVYLLIHGTADDNVHVFQOQAISKALV 60
QY 699 DVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSL 739
Db 61 NQAVDFQAMWYSDQNHL-SGLSTNHLTYTHMTFLKQCFSL 100
```

## RESULT 7

```
US-11-176-951-5
; Sequence 5, Application US/11176951
; Publication No. US20060024313A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, XIN
; APPLICANT: CHEN, YUAN-SHOU
; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
; TITLE OF INVENTION: PROLYL DIPEPTIDASES
; FILE REFERENCE: 08842.0019
; CURRENT APPLICATION NUMBER: US/11/176,951
; CURRENT FILING DATE: 2005-07-06
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952
; PRIOR FILING DATE: 2004-07-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 5
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-5
```

```
Query Match 7.2%; Score 291; DB 7; Length 54;
Best Local Similarity 100.0%; Pred. No. 2.3e-16;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 687 FQOQAISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP 740
Db 1 FQOQAISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP 54
```

## RESULT 8

```
US-11-179-977-1
; Sequence 1, Application US/11179977
; Publication No. US20050249789A1
; GENERAL INFORMATION:
; APPLICANT: Genencor International, Inc.
; TITLE OF INVENTION: Alpha/Beta Hydrolase-Fold Enzymes
; FILE REFERENCE: GC511-PCT
; CURRENT APPLICATION NUMBER: US/11/179,977
; CURRENT FILING DATE: 2005-07-12
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Bacillus
US-11-179-977-1
```



```
; PRIOR FILING DATE: 2004-07-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 6
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-6

Query Match      3.6%; Score 145; DB 7; Length 27;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 208 PLIEYSFYSDLSQYPKTVRVYPKAG 234
Db 1 PLIEYSFYSDLSQYPKTVRVYPKAG 27

RESULT 12
US-11-176-951-13
; Sequence 13, Application US/11176951
; Publication No. US20060024313A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, XIN
; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
; TITLE OF INVENTION: PROLYL DIPEPTIDASES
; FILE REFERENCE: 08842.0019
; CURRENT APPLICATION NUMBER: US/11/176,951
; PRIOR FILING DATE: 2005-07-06
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 13
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-13

Query Match      3.6%; Score 145; DB 7; Length 27;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 208 PLIEYSFYSDLSQYPKTVRVYPKAG 234
Db 1 PLIEYSFYSDLSQYPKTVRVYPKAG 27

RESULT 13
US-11-176-951-10
; Sequence 10, Application US/11176951
; Publication No. US20060024313A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, XIN
; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
; TITLE OF INVENTION: PROLYL DIPEPTIDASES
; FILE REFERENCE: 08842.0019
; CURRENT APPLICATION NUMBER: US/11/176,951
; CURRENT FILING DATE: 2005-07-06
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 10
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
US-11-176-951-10

Query Match      3.6%; Score 145; DB 7; Length 109;
Best Local Similarity 33.7%; Pred. No. 0.00013;
Matches 33; Conservative 17; Mismatches 44; Indels 4; Gaps 2;

QY 640 YTERVMGLPTPEDNLDHYRNSTVMSRAENF--KQVEYLLIHGTADDNVHFQSOAQISKAL 697
Db 2 YTERVMG--HPDQNEGGYILGSVAMQAEKFPSEPNRLILLHGFLENVHFAHTSILLSFL 59

QY 698 VDVGVDFQAMMYTDEDHGFIASSTAHOHIYTHMSHFQ 735
Db 60 VRAGKPYDLQIYPQSRHSIRVPESGEGHYELHLLHYLQE 97

RESULT 14
US-11-176-951-11
; Sequence 11, Application US/11176951
; Publication No. US20060024313A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, XIN
; TITLE OF INVENTION: AGENTS THAT DISRUPT DIMER FORMATION IN DPP-IV FAMILY OF
; TITLE OF INVENTION: PROLYL DIPEPTIDASES
; FILE REFERENCE: 08842.0019
; CURRENT APPLICATION NUMBER: US/11/176,951
; CURRENT FILING DATE: 2005-07-06
; PRIOR APPLICATION NUMBER: 60/586,095
; PRIOR FILING DATE: 2004-07-06
; PRIOR APPLICATION NUMBER: 60/585,952
; PRIOR FILING DATE: 2004-07-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 11
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-176-951-11

Query Match      3.2%; Score 130; DB 7; Length 99;
Best Local Similarity 30.6%; Pred. No. 0.0017;
Matches 30; Conservative 18; Mismatches 46; Indels 4; Gaps 2;

QY 640 YTERVMGLPTPEDNLDHYRNSTVMSRAENF--KQVEYLLIHGTADDNVHFQSOAQISKAL 697
Db 2 YTERVMDV--PENNOGHYEAGSVLHVEXLPNEPNRLILLHGFLENVHFAHTSILLSFL 59

QY 698 VDVGVDFQAMMYTDEDHGFIASSTAHOHIYTHMSHFQ 735
Db 60 IRAGKPYQLQIYPNERHSIRCPESGEGHYEVTLHLFLQE 97

RESULT 15
US-10-454-437-118
; Sequence 118, Application US/10454437
; Publication No. US20050277115A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Krogger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberkauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN HOMEOSTASIS AND ADAPTATION
; FILE REFERENCE: BGI-128CPCN
; CURRENT APPLICATION NUMBER: US/10/454,437
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
```

```
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 118
; LENGTH: 668
; TYPE: PRF
; ORGANISM: Corynebacterium glutamicum
US-10-454-437-118

Query Match      3.1%; Score 124; DB 6; Length 668;
Best Local Similarity 19.5%; Pred. No. 0.065;
Matches 101; Conservative 68; Mismatches 180; Indels 170; Gaps 23;

Qy 321 EMSTTGWGRPRSEPHF-----TLDGNSFYKIISNEGYRHICYFOID-KKDCFTFIT 372
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
212 DVATHW-----RDTFGERTFVSRLDPYNSETSLTEGG-----LVKLDVPTDCDIV 262
Qy 373 KGTWEVIGIEALTSDYLYIS--NEYKGMPCGRNLYKQLSDYTKVTKLSCELNPERCQY 430
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
263 KKQW-----IFVSPRTDFAGIPAG-GLGVLLKX----- 290
Qy 431 YSVSFSKEAKYQLRCSPGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVOM----- 483
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
291 -----FLEGGGRDFQ-----PVFTPTSTSLQSLATTK-NFLVLTLLNNVSTEIVTPL 337
Qy 484 --PSKLDLFIILNE-----TKFWYQMI-----LPPHFDKSKKYP 515
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
338 NDPTEHEHIDLPEHVTAHVVATSPLDGDEIHWQASFTPEPTLLRAELPGALEAVKAP 397
Qy 516 LLDVYAGPCSKADTVFRLNWAYTLAST-----ENIIIVASFDG----- 554
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
398 LQF-----ENAGQETQHWATSADQTKIPYFITGAFEEBPQNTLVHAYGFEVSLTP 449
Qy 555 -----RSGYQGDKIMHAINRRLGTPEV-EDQIEAARQPSKMG 591
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
450 SHSPTRGIAMLEKGYFVEANLRGGSGFGPE-WHSQATKLRMKVWEDHRAVLADLVERG 508
Qy 592 FVDNKRITAIWGSYGVVTSMLVSGSGVFKCGIAPVPSRWYVY-----DSVYTERYMG 646
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
509 YATPEQIAIRGGSGGLTSGALTQYPAFGAAVVQVPLADMLRYHTWSAGASWMAEYGN 568
Qy 647 LFTPEDN--LDHYRNSTVMGRAENFKQVYLLIHGTADDNNVHFOQSAISKALVDVGVD 704
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
569 PDPPEERAVIEQVSPQAVGVGVEKRIYPPALVTTSTRDRVHPAHARLFAQALLDAG--- 625
Qy 705 QAM-WYTDDEHGIASSTAQH-----IYTHMSHPKQ 735
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
626 QAVDYENTEGGHAGAADNKQAFVESLIYTWIEKTLTLDQ 664
```

Search completed: February 15, 2006, 13:29:41  
Job time : 19 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2006, 13:25:59 ; Search time 168 Seconds  
(without alignments)  
1840.439 Million cell updates/sec

Title: US-10-659-055-3

Perfect score: 4020

Sequence: 1 ADPGSHHHHSRKYTLT.....AQHIYTHMSHFQKCSLP 740

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA\_Main:  
1: /cgm2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgm2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
3: /cgm2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
4: /cgm2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
5: /cgm2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
6: /cgm2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3939	98.0	766	3	US-09-265-606-3
2	3939	98.0	766	3	US-09-993-959-1
3	3939	98.0	766	4	US-10-295-027-590
4	3939	98.0	766	5	US-10-952-459-18
5	3939	98.0	766	5	US-10-631-467-565
6	3933	97.8	766	4	US-10-002-593-6
7	3933	97.8	766	4	US-10-165-603-7
8	3933	97.8	766	4	US-10-423-714-6
9	3933	97.8	766	4	US-10-295-027-922
10	3933	97.8	766	4	US-10-794-899-41
11	3933	97.8	766	6	US-11-041-674-6
12	3929	97.7	766	5	US-10-476-264-147
13	3929	97.7	766	5	US-10-770-712-1
14	3466.5	84.7	767	5	US-10-770-712-2
15	3395.5	84.5	767	4	US-10-165-603-4
16	3395.5	84.5	767	4	US-10-794-899-38
17	3390	84.3	760	5	US-10-631-467-1390
18	3374	83.9	760	5	US-10-770-712-3
19	2175	54.1	760	5	US-10-723-860-4171
20	2168	53.9	760	3	US-09-265-606-2
21	2168	53.9	760	4	US-10-177-293-136
22	2168	53.9	760	4	US-10-301-822-55
23	2168	53.9	760	4	US-10-884-070A-13
24	1289.5	32.1	504	4	US-10-072-012-863
25	1223	30.4	746	5	US-10-476-264-69
26	1223	30.4	746	5	US-10-476-264-70
27	1223	30.4	746	5	US-10-476-264-72

28	1223	30.4	789	5	US-10-476-264-71	Sequence 71, Appl
29	1223	30.4	796	3	US-09-976-674-5	Sequence 5, Appl
30	1223	30.4	796	3	US-09-870-133-2	Sequence 2, Appl
31	1223	30.4	796	4	US-10-160-501-5	Sequence 5, Appl
32	1223	30.4	796	5	US-10-982-512-5	Sequence 5, Appl
33	1223	30.4	796	5	US-10-433-757-2	Sequence 2, Appl
34	1223	30.4	796	5	US-10-476-264-68	Sequence 68, Appl
35	1223	30.4	796	5	US-10-476-264-110	Sequence 110, App
36	1223	30.4	796	5	US-10-476-264-110	Sequence 106, App
37	1223	30.4	796	5	US-10-476-264-142	Sequence 142, App
38	1223	30.4	796	5	US-10-476-264-146	Sequence 146, App
39	1223	30.4	797	5	US-10-476-264-36	Sequence 36, Appl
40	1221	30.4	228	5	US-10-476-264-114	Sequence 114, App
41	1217	30.3	798	4	US-10-210-130-120	Sequence 120, App
42	1198	29.8	743	4	US-10-363-937-4	Sequence 4, Appl
43	1198	29.8	743	6	US-11-167-040-41	Sequence 4, Appl
44	1196	29.8	706	3	US-09-976-674-41	Sequence 41, Appl
45	1196	29.8	706	5	US-10-982-512-41	Sequence 41, Appl

ALIGNMENTS

RESULT 1  
US-09-265-606-3  
; Sequence 3, Application US/09265606  
; Patent No. US20020034789A1  
; GENERAL INFORMATION:  
; APPLICANT: Zimmermann, Rainer; Park, John E.;  
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.  
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN  
; TITLE OF INVENTION: ALPHA, AND USES THEREOF  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felfe & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/265,606  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/619,280  
FILING DATE: 18-MARCH-1996  
APPLICATION NUMBER: 08/230,491  
FILING DATE: 20-APRIL-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. US20020034789A1man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5330.1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 766 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-09-265-606-3

Query Match 98.0%; Score 3939; DB 3; Length 766;  
Best Local Similarity 100.0%; Pred. No. 3.7e-306;  
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 13 SRKTYTLTDYLNKTYRLKLYSLRWISDHEYLKQENILVFNABYGNSSVFLENSTFDEF 72

```

Db      39 SRKTYTLTDYLNKNTYRLKLYSLRWISDHEYLKQENNLVFNABVGNSSVFLNSTDDEF 98
      73 GHSINDYSIPDQGFILLEYVVKQWRHSYTTASDYIDLKQRLITEERIIPNNTQVWTWS 132
      99 GHSINDYSIPDQGFILLEYVVKQWRHSYTTASDYIDLKQRLITEERIIPNNTQVWTWS 158
      133 PVGHKLAYVWNNDIYVKIEPNLPSYRIITWTGKEDIYNGITDWTVEEVEFSAYSALWWS 192
      159 PVGHKLAYVWNNDIYVKIEPNLPSYRIITWTGKEDIYNGITDWTVEEVEFSAYSALWWS 218
      193 NGTEFLAYAQFNDTEVPLIEYFYSDESLOYPKTVRPVYPKAGAVNPTVKPFVNTDLS 252
      219 NGTEFLAYAQFNDTEVPLIEYFYSDESLOYPKTVRPVYPKAGAVNPTVKPFVNTDLS 278
      253 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLOWLRRIQNSVMDICDYDESSGRWN 312
      279 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLOWLRRIQNSVMDICDYDESSGRWN 338
      313 CLVARQHIEMSTGWGRFRPSEPHTLDGNSFYKIIISNEEGYRHCYFQIDKKDCTFT 372
      339 CLVARQHIEMSTGWGRFRPSEPHTLDGNSFYKIIISNEEGYRHCYFQIDKKDCTFT 398
      373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS 432
      399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS 458
      433 VSFSEAKYQYLRCSGPGPLVTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFTI 492
      459 VSFSEAKYQYLRCSGPGPLVTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFTI 518
      493 LNETKFWYQMLPPHFDKSKYPLLDVYAGCSQKADTVFRLNWTYLASTENIIVASF 552
      519 LNETKFWYQMLPPHFDKSKYPLLDVYAGCSQKADTVFRLNWTYLASTENIIVASF 578
      553 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWNGSYGGYVTSM 612
      579 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWNGSYGGYVTSM 638
      613 VLGSQGVFKCGIAPVSRWEYDSVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQV 672
      639 VLGSQGVFKCGIAPVSRWEYDSVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQV 698
      673 EYLLIHGTADDNVHFOQSAQISKALVDVGVDFQAMWYTTDEDHGIASTAHQHIYTHMSHF 732
      699 EYLLIHGTADDNVHFOQSAQISKALVDVGVDFQAMWYTTDEDHGIASTAHQHIYTHMSHF 758
      733 IKQCFSLP 740
      759 IKQCFSLP 766

```

RESULT 2

```

US-09-993-959-1
; Sequence 1, Application US/0993959
; Publication No. US20030165489A1
; GENERAL INFORMATION:
; APPLICANT: Grouzmann, Eric
; APPLICANT: Lacroix, Jean-Silvain
; APPLICANT: Monod, Michel
; TITLE OF INVENTION: Method of Treating Rhinitis and Sinusitis
; FILE REFERENCE: 8195/276823
; CURRENT APPLICATION NUMBER: US/09/993,959
; CURRENT FILING DATE: 2001-11-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-993-959-1

```

Query Match 98.0%; Score 3939; DB 3; Length 766;

```

      13 SRKTYTLTDYLNKNTYRLKLYSLRWISDHEYLKQENNLVFNABVGNSSVFLNSTDDEF 72
      39 SRKTYTLTDYLNKNTYRLKLYSLRWISDHEYLKQENNLVFNABVGNSSVFLNSTDDEF 98
      73 GHSINDYSIPDQGFILLEYVVKQWRHSYTTASDYIDLKQRLITEERIIPNNTQVWTWS 132
      99 GHSINDYSIPDQGFILLEYVVKQWRHSYTTASDYIDLKQRLITEERIIPNNTQVWTWS 158
      133 PVGHKLAYVWNNDIYVKIEPNLPSYRIITWTGKEDIYNGITDWTVEEVEFSAYSALWWS 192
      159 PVGHKLAYVWNNDIYVKIEPNLPSYRIITWTGKEDIYNGITDWTVEEVEFSAYSALWWS 218
      193 NGTEFLAYAQFNDTEVPLIEYFYSDESLOYPKTVRPVYPKAGAVNPTVKPFVNTDLS 252
      219 NGTEFLAYAQFNDTEVPLIEYFYSDESLOYPKTVRPVYPKAGAVNPTVKPFVNTDLS 278
      253 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLOWLRRIQNSVMDICDYDESSGRWN 312
      279 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLOWLRRIQNSVMDICDYDESSGRWN 338
      313 CLVARQHIEMSTGWGRFRPSEPHTLDGNSFYKIIISNEEGYRHCYFQIDKKDCTFT 372
      339 CLVARQHIEMSTGWGRFRPSEPHTLDGNSFYKIIISNEEGYRHCYFQIDKKDCTFT 398
      373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS 432
      399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS 458
      433 VSFSEAKYQYLRCSGPGPLVTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFTI 492
      459 VSFSEAKYQYLRCSGPGPLVTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFTI 518
      493 LNETKFWYQMLPPHFDKSKYPLLDVYAGCSQKADTVFRLNWTYLASTENIIVASF 552
      519 LNETKFWYQMLPPHFDKSKYPLLDVYAGCSQKADTVFRLNWTYLASTENIIVASF 578
      553 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWNGSYGGYVTSM 612
      579 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWNGSYGGYVTSM 638
      613 VLGSQGVFKCGIAPVSRWEYDSVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQV 672
      639 VLGSQGVFKCGIAPVSRWEYDSVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQV 698
      673 EYLLIHGTADDNVHFOQSAQISKALVDVGVDFQAMWYTTDEDHGIASTAHQHIYTHMSHF 732
      699 EYLLIHGTADDNVHFOQSAQISKALVDVGVDFQAMWYTTDEDHGIASTAHQHIYTHMSHF 758
      733 IKQCFSLP 740
      759 IKQCFSLP 766

```

RESULT 3

```

US-10-295-027-590
; Sequence 590, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; Methods of Screening for Modulators of Cancer

```

FILE REFERENCE: 018501-012500US  
; CURRENT APPLICATION NUMBER: US/10/295,027  
; CURRENT FILING DATE: 2002-11-13  
; PRIOR APPLICATION NUMBER: US 09/663,733  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: US 60/350,666  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/335,394  
; PRIOR FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/332,464  
; PRIOR FILING DATE: 2001-11-21  
; PRIOR APPLICATION NUMBER: US 60/334,393  
; PRIOR FILING DATE: 2001-11-29  
; PRIOR APPLICATION NUMBER: US 60/340,376  
; PRIOR FILING DATE: 2001-12-14  
; PRIOR APPLICATION NUMBER: US 60/347,211  
; PRIOR FILING DATE: 2002-01-08  
; PRIOR APPLICATION NUMBER: US 60/347,349  
; PRIOR FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 60/355,250  
; PRIOR FILING DATE: 2002-02-08  
; PRIOR APPLICATION NUMBER: US 60/356,714  
; PRIOR FILING DATE: 2002-02-13  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1386  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 590  
; LENGTH: 766  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-295-027-590

Query Match 98.0%; Score 3939; DB 4; Length 766;  
Best Local Similarity 100.0%; Pred. No. 3.7e-306;  
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 SRKTYTLDYLNKNTYRLKLSLRWISDHEYLKQENNLVFNABYGNSSVFLNSTPDEF 72  
Db 39 SRKTYTLDYLNKNTYRLKLSLRWISDHEYLKQENNLVFNABYGNSSVFLNSTPDEF 98  
Qy 73 GHSINDYSISPDGQFILLENNYVQWRHSYTSYDIYDLNKRQLITERIPNNTQWWTWS 132  
Db 99 GHSINDYSISPDGQFILLENNYVQWRHSYTSYDIYDLNKRQLITERIPNNTQWWTWS 158  
Qy 133 PVGHKLAVVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWSP 192  
Db 159 PVGHKLAVVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWSP 218  
Qy 193 NGTFLAYAFNDTEVPLIEYSFYSDLSQYPKTVRVPYKAGAVNPTVKPFVNTDSLSS 252  
Db 219 NGTFLAYAFNDTEVPLIEYSFYSDLSQYPKTVRVPYKAGAVNPTVKPFVNTDSLSS 278  
Qy 253 VTNATSIQITAPASMLGDHYLCDVTWATQERISLQWLRRIQNTSVMDICDYDESSGRWN 312  
Db 279 VTNATSIQITAPASMLGDHYLCDVTWATQERISLQWLRRIQNTSVMDICDYDESSGRWN 338  
Qy 313 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEBGRYHICVFQIDKKDCTFIT 372  
Db 339 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEBGRYHICVFQIDKKDCTFIT 398  
Qy 373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCISCELNPERCQYIS 432  
Db 399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCISCELNPERCQYIS 458  
Qy 433 VSPSKEAKYQLRCSGGLPLTYTLHSSVNDKGLRVLSDNSALDKMLQNVQMPSKKLDFFII 492  
Db 459 VSPSKEAKYQLRCSGGLPLTYTLHSSVNDKGLRVLSDNSALDKMLQNVQMPSKKLDFFII 518  
Qy 493 LNETKFWQMLPPHFDKSKYKYPILLDVIYAGPCSKADTVFRLNWTATYLASTENIIVASF 552  
Db 519 LNETKFWQMLPPHFDKSKYKYPILLDVIYAGPCSKADTVFRLNWTATYLASTENIIVASF 578  
Qy 553 DGRSGYQGDKIMHAINRRLGTFFVEVDQIEAARQFSKMGFVDNKRRIALWGSYGYVTSM 612

Db 579 DGRSGYQGDKIMHAINRRLGTFFVEVDQIEAARQFSKMGFVDNKRRIALWGSYGYVTSM 638  
Qy 613 VLGSAGVFKCGIAVAPVSRWEYDVSYYTERYMGLPTEPDNLNHYRNSTVMSRAENPKOV 672  
Db 639 VLGSAGVFKCGIAVAPVSRWEYDVSYYTERYMGLPTEPDNLNHYRNSTVMSRAENPKOV 698  
Qy 673 EYLLIHGTADDNVHFQOQAISKALVDVGVDFQAWWYTDDEHGIASSTAHOHIYTHMSHF 732  
Db 699 EYLLIHGTADDNVHFQOQAISKALVDVGVDFQAWWYTDDEHGIASSTAHOHIYTHMSHF 758  
Qy 733 IKQCFSLP 740  
Db 759 IKQCFSLP 766

RESULT 4  
US-10-952-459-18  
; Sequence 18, Application US/10952459  
; Publication No. US20050074805A1  
; GENERAL INFORMATION:  
; APPLICANT: Kochan, Jarema Peter  
; APPLICANT: Martin, Mitchell Lee  
; APPLICANT: Robinski, James Andrew  
; TITLE OF INVENTION: Specific Markers for Diabetes  
; FILE REFERENCE: 21270US1  
; CURRENT APPLICATION NUMBER: US/10/952,459  
; CURRENT FILING DATE: 2004-09-28  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 18  
; LENGTH: 766  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-952-459-18

Query Match 98.0%; Score 3939; DB 5; Length 766;  
Best Local Similarity 100.0%; Pred. No. 3.7e-306;  
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 SRKTYTLDYLNKNTYRLKLSLRWISDHEYLKQENNLVFNABYGNSSVFLNSTPDEF 72  
Db 39 SRKTYTLDYLNKNTYRLKLSLRWISDHEYLKQENNLVFNABYGNSSVFLNSTPDEF 98  
Qy 73 GHSINDYSISPDGQFILLENNYVQWRHSYTSYDIYDLNKRQLITERIPNNTQWWTWS 132  
Db 99 GHSINDYSISPDGQFILLENNYVQWRHSYTSYDIYDLNKRQLITERIPNNTQWWTWS 158  
Qy 133 PVGHKLAVVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWSP 192  
Db 159 PVGHKLAVVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWSP 218  
Qy 193 NGTFLAYAFNDTEVPLIEYSFYSDLSQYPKTVRVPYKAGAVNPTVKPFVNTDSLSS 252  
Db 219 NGTFLAYAFNDTEVPLIEYSFYSDLSQYPKTVRVPYKAGAVNPTVKPFVNTDSLSS 278  
Qy 253 VTNATSIQITAPASMLGDHYLCDVTWATQERISLQWLRRIQNTSVMDICDYDESSGRWN 312  
Db 279 VTNATSIQITAPASMLGDHYLCDVTWATQERISLQWLRRIQNTSVMDICDYDESSGRWN 338  
Qy 313 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEBGRYHICVFQIDKKDCTFIT 372  
Db 339 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEBGRYHICVFQIDKKDCTFIT 398  
Qy 373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCISCELNPERCQYIS 432  
Db 399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCISCELNPERCQYIS 458  
Qy 433 VSPSKEAKYQLRCSGGLPLTYTLHSSVNDKGLRVLSDNSALDKMLQNVQMPSKKLDFFII 492  
Db 459 VSPSKEAKYQLRCSGGLPLTYTLHSSVNDKGLRVLSDNSALDKMLQNVQMPSKKLDFFII 518  
Qy 493 LNETKFWQMLPPHFDKSKYKYPILLDVIYAGPCSKADTVFRLNWTATYLASTENIIVASF 552

[illegible]

RESULT 5  
 US-10-631-467-565  
 ; Sequence 565, Application US/10631467  
 ; Publication No. US20050208496A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Genox Research Inc.  
 ; TITLE OF INVENTION: Method for testing for broncheal asthma, or chronic obstructive p  
 ; TITLE OF INVENTION: disease  
 ; FILE REFERENCE: 3462.1005-000  
 ; CURRENT APPLICATION NUMBER: US/10/631.467  
 ; CURRENT FILING DATE: 2003-07-31  
 ; PRIOR APPLICATION NUMBER: JP 2003-077212  
 ; PRIOR FILING DATE: 2003-03-20  
 ; PRIOR APPLICATION NUMBER: JP 2002-229312  
 ; PRIOR FILING DATE: 2002-08-06  
 ; NUMBER OF SEQ ID NOS: 2086  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 565  
 ; LENGTH: 766  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-631-467-565

Query Match 98.0%; Score 3939; DB 5; Length 766;  
Best Local Similarity 100.0%; Pred. No. 3.7e-306;  
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	13	SRKYTYLLTDYLYKNYTRKLYLSRWISDHEHYLYKQENNLVFNABVCNSVFLNSTFDFE	72
Db	39	SRKYTYLLTDYLYKNYTRKLYLSRWISDHEHYLYKQENNLVFNABVCNSVFLNSTFDFE	98
Qy	73	GHSINDYSGISPDGQFILLEYNVVKQWRISYTHSYDYLDLKKQLITEERI PNNTQWTVWS	132
Db	99	GHSINDYSGISPDGQFILLEYNVVKQWRISYTHSYDYLDLKKQLITEERI PNNTQWTVWS	158
Qy	133	PVGHKLAYVMNNDIYVVKIEPNLPSYRIITWTGKEDIYNGITDQWVYEEBVFVSALWWS	192
Db	159	PVGHKLAYVMNNDIYVVKIEPNLPSYRIITWTGKEDIYNGITDQWVYEEBVFVSALWWS	218
Qy	193	NGTFLAYAQFNDETVPLLEYSPYSDESIYQPKTVRPVYPKAGAVNPTVAFVYVNTDLS	252
Db	219	NGTFLAYAQFNDETVPLLEYSPYSDESIYQPKTVRPVYPKAGAVNPTVAFVYVNTDLS	278
Qy	253	VTNATSIQITAPASMLIGDHYLCDVTWATQERISLQWLRIQNSYVMDICYDESSGRWN	312
Db	279	VTNATSIQITAPASMLIGDHYLCDVTWATQERISLQWLRIQNSYVMDICYDESSGRWN	338
Qy	313	CLVARQHLTMSYTGWGRPRPSEPHFTLDGNSFYKLIISNEEGYRHYCYFQIDKKOCTFT	372
Db	339	CLVARQHLTMSYTGWGRPRPSEPHFTLDGNSFYKLIISNEEGYRHYCYFQIDKKOCTFT	398
Qy	373	KGTWEVIGIEALTSYLYYISNEYKMPGGRNLYKIQLSDYTKVTCLSCELNPERQYYS	432

[illegible]

```

RESULT 6
US-10-002-593-6
; Sequence 6, Application US/10002593
; Publication No. US20020137120A1
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Brown, Nancy J.
; TITLE OF INVENTION: BIOLOGICAL MARKERS FOR ANGIOTENSIN CONVERTIN
; TITLE OF INVENTION: INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA
; FILE REFERENCE: Atty Docket No. US20020137120A1 1242/48/2
; CURRENT APPLICATION NUMBER: US/10/002,593
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/244,524
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-002-593-6

```

Query Match 97.8%; Score 3933; DB 4; Length 766;  
Best Local Similarity 99.9%; Pred. No. 1.le-305;

Qy	13	SRKTYTLTDYLYKNTYRLKLYSLRWISDRHEYLYKQENNILVFNABEYVNSVFLNSTFDEF	772
Db	39	SRKTYTLTDYLYKNTYRLKLYSLRWISDRHEYLYKQENNILVFNABEYVNSVFLNSTFDEF	98
Qy	73	GHSINDYSISPDGQFILLEVNVVQWRHSYATSYDIYDLNKKQLITEERIPNNTQWVTWS	132
Db	99	GHSINDYSISPDGQFILLEVNVVQWRHSYATSYDIYDLNKKQLITEERIPNNTQWVTWS	158
Qy	133	PVGKLAAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDQWYEEVFSAYSALWWS	192
Db	159	PVGKLAAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDQWYEEVFSAYSALWWS	218
Qy	193	NGTFLLAQAQNDTEVPLEIYSFYSDSLQYPKTVRPVYPKAGAVNPTVKFFVNTDLSLS	252
Db	219	NGTFLLAQAQNDTEVPLEIYSFYSDSLQYPKTVRPVYPKAGAVNPTVKFFVNTDLSLS	278
Qy	253	VTNATSIQITAPASMLIGDHYLDCVTWATQBRISLQWLRRINQYSVMDICDYDESSGRWN	312
Db	279	VTNATSIQITAPASMLIGDHYLDCVTWATQBRISLQWLRRINQYSVMDICDYDESSGRWN	338

QY 313 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNBERGYRHCYFQIDKKDCTFIT 372  
DB 339 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNBERGYRHCYFQIDKKDCTFIT 398  
QY 373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYVS 432  
DB 399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYVS 458  
QY 433 VPSKBAKYYQLRCSGPGPLPYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPSPKCLDFII 492  
DB 459 VPSKBAKYYQLRCSGPGPLPYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPSPKCLDFII 518  
QY 493 LNETKFWYQMLPPLPHFDKSKYKYPILLDVIYAGPCSKADTVFRLNWTATYLASTENIIVASF 552  
DB 519 LNETKFWYQMLPPLPHFDKSKYKYPILLDVIYAGPCSKADTVFRLNWTATYLASTENIIVASF 578  
QY 553 DGRSGYQGDKIMHAINRRLGTPEVEDQIEAARQFSKMGFVDNKRKRIATWGSYGGYVTSM 612  
DB 579 DGRSGYQGDKIMHAINRRLGTPEVEDQIEAARQFSKMGFVDNKRKRIATWGSYGGYVTSM 638  
QY 613 VLGSAGVFKGCIAPVPSRWEYDVSVYTERYMGSLPTPEDNLDHYRNSTVMSRAENFKQV 672  
DB 639 VLGSAGVFKGCIAPVPSRWEYDVSVYTERYMGSLPTPEDNLDHYRNSTVMSRAENFKQV 698  
QY 673 EYLLIHGTADDNVHFQOQAISKALVDVGVDFQAMWYTDDEHGIIASSTAHQHIYTHMSHF 732  
DB 699 EYLLIHGTADDNVHFQOQAISKALVDVGVDFQAMWYTDDEHGIIASSTAHQHIYTHMSHF 758  
QY 733 IKQCFSLP 740  
DB 759 IKQCFSLP 766

RESULT 7  
US-10-165-603-7  
; Sequence 7, Application US/10165603  
; Publication No. US20030021792A1  
; GENERAL INFORMATION:  
; APPLICANT: Roben, Paul W.  
; TITLE OF INVENTION: TISSUE-SPECIFIC ENDOTHELIAL MEMBRANE  
; TITLE OF INVENTION: PROTEINS  
; FILE REFERENCE: TPTECH.001A  
; CURRENT APPLICATION NUMBER: US/10/165.603  
; CURRENT FILING DATE: 2002-06-07  
; PRIOR APPLICATION NUMBER: 60/297,021  
; PRIOR FILING DATE: 2001-06-08  
; PRIOR APPLICATION NUMBER: 60/305,117  
; PRIOR FILING DATE: 2001-07-12  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 766  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-165-603-7

Query Match 97.8%; Score 3933; DB 4; Length 766;  
Best Local Similarity 99.9%; Pred. No. 1.1e-305;  
Matches 727; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 SRKTYTLTDVLYKNTYRLKLSLRWISDHEVLYKQENNILVFNABYGNSSVPLENSTDFEF 72  
DB 39 SRKTYTLTDVLYKNTYRLKLSLRWISDHEVLYKQENNILVFNABYGNSSVPLENSTDFEF 98  
QY 73 GHSINDYSISPDGQFILLLENNYKQWRHSYTPASDYIDLANKRQLITEERIPNNQTQWTS 132  
DB 99 GHSINDYSISPDGQFILLLENNYKQWRHSYTPASDYIDLANKRQLITEERIPNNQTQWTS 158  
QY 133 PVGHKLAYWNNNDIYVKEPNLPSYRITWTCKEDIIYNGITDWTYEEVFSAYSALWSP 192  
DB 159 PVGHKLAYWNNNDIYVKEPNLPSYRITWTCKEDIIYNGITDWTYEEVFSAYSALWSP 218

QY 193 NGTFLAVAFNDTEVPLIEYSFYSDLESQYPKTVRVPYKAGAVNPTVKFFVYVNTDSLSS 252  
DB 219 NGTFLAVAFNDTEVPLIEYSFYSDLESQYPKTVRVPYKAGAVNPTVKFFVYVNTDSLSS 278  
QY 253 VTNATSIQITAPASMLIGDHYLCDVTWATQBRISLOWLRRIQNTYSVMDICDYDESSGRWN 312  
DB 279 VTNATSIQITAPASMLIGDHYLCDVTWATQBRISLOWLRRIQNTYSVMDICDYDESSGRWN 338  
QY 313 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNBERGYRHCYFQIDKKDCTFIT 372  
DB 339 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNBERGYRHCYFQIDKKDCTFIT 398  
QY 373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYVS 432  
DB 399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYVS 458  
QY 433 VPSKBAKYYQLRCSGPGPLPYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPSPKCLDFII 492  
DB 459 VPSKBAKYYQLRCSGPGPLPYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPSPKCLDFII 518  
QY 493 LNETKFWYQMLPPLPHFDKSKYKYPILLDVIYAGPCSKADTVFRLNWTATYLASTENIIVASF 552  
DB 519 LNETKFWYQMLPPLPHFDKSKYKYPILLDVIYAGPCSKADTVFRLNWTATYLASTENIIVASF 578  
QY 553 DGRSGYQGDKIMHAINRRLGTPEVEDQIEAARQFSKMGFVDNKRKRIATWGSYGGYVTSM 612  
DB 579 DGRSGYQGDKIMHAINRRLGTPEVEDQIEAARQFSKMGFVDNKRKRIATWGSYGGYVTSM 638  
QY 613 VLGSAGVFKGCIAPVPSRWEYDVSVYTERYMGSLPTPEDNLDHYRNSTVMSRAENFKQV 672  
DB 639 VLGSAGVFKGCIAPVPSRWEYDVSVYTERYMGSLPTPEDNLDHYRNSTVMSRAENFKQV 698  
QY 673 EYLLIHGTADDNVHFQOQAISKALVDVGVDFQAMWYTDDEHGIIASSTAHQHIYTHMSHF 732  
DB 699 EYLLIHGTADDNVHFQOQAISKALVDVGVDFQAMWYTDDEHGIIASSTAHQHIYTHMSHF 758  
QY 733 IKQCFSLP 740  
DB 759 IKQCFSLP 766

RESULT 8  
US-10-423-714-6  
; Sequence 6, Application US/10423714  
; Publication No. US20030180828A1  
; GENERAL INFORMATION:  
; APPLICANT: Vanderbilt University  
; APPLICANT: Brown, Nancy J.  
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTING ENZYME  
; TITLE OF INVENTION: INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA  
; FILE REFERENCE: Acty Docket No. US20030180828A1 1242/48/2  
; CURRENT APPLICATION NUMBER: US/10/423,714  
; CURRENT FILING DATE: 2003-04-25  
; PRIOR APPLICATION NUMBER: 60/244,524  
; PRIOR FILING DATE: 2000-10-31  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 766  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-423-714-6

Query Match 97.8%; Score 3933; DB 4; Length 766;  
Best Local Similarity 99.9%; Pred. No. 1.1e-305;  
Matches 727; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 SRKTYTLTDVLYKNTYRLKLSLRWISDHEVLYKQENNILVFNABYGNSSVPLENSTDFEF 72  
DB 39 SRKTYTLTDVLYKNTYRLKLSLRWISDHEVLYKQENNILVFNABYGNSSVPLENSTDFEF 98  
QY 73 GHSINDYSISPDGQFILLLENNYKQWRHSYTPASDYIDLANKRQLITEERIPNNQTQWTS 132

Db 99 GHSINDYSIPDQFILLLEYNVVKQWRHSYASDYIDLKQRLTEERIPNNTQWVTWS 158  
Qy 133 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFSAYSALWSP 192  
Db 159 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFSAYSALWSP 218  
Qy 193 NGTEFLAYAQFNDTEVPLIEYSFYDESLOYPKTVRVPYKAGAVNPTVKFFVNTDLSLSS 252  
Db 219 NGTEFLAYAQFNDTEVPLIEYSFYDESLOYPKTVRVPYKAGAVNPTVKFFVNTDLSLSS 278  
Qy 253 VTNATSIQITAPASMLIGDHYLCVDTWATQERISLQWLRRIQNYSVMDICDYDESSGRWN 312  
Db 279 VTNATSIQITAPASMLIGDHYLCVDTWATQERISLQWLRRIQNYSVMDICDYDESSGRWN 338  
Qy 313 CLVARQHIEMSTTGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCCTFIT 372  
Db 339 CLVARQHIEMSTTGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCCTFIT 398  
Qy 373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCCLSELNPERCOYYS 432  
Db 399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCCLSELNPERCOYYS 458  
Qy 433 VSFSEAKYQYLRCSGPGPLLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPSSKKLDFII 492  
Db 459 VSFSEAKYQYLRCSGPGPLLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPSSKKLDFII 518  
Qy 493 LNETKFWYQMLPPHFDKSKYPLLLDVYAGPSCOKADTVFRLNWTYLASTENIIVASF 552  
Db 519 LNETKFWYQMLPPHFDKSKYPLLLDVYAGPSCOKADTVFRLNWTYLASTENIIVASF 578  
Qy 553 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGVVTSM 612  
Db 579 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGVVTSM 638  
Qy 613 VLGSYGVPKCGIAVAPVSRWEYDVSVYTERYMGLTPTPEDNLDHYNSTVMSRAENFKQV 672  
Db 639 VLGSYGVPKCGIAVAPVSRWEYDVSVYTERYMGLTPTPEDNLDHYNSTVMSRAENFKQV 698  
Qy 673 EYLLIHGTADDNVHFOQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHF 732  
Db 699 EYLLIHGTADDNVHFOQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHF 758  
Qy 733 IKQCPSLP 740  
Db 759 IKQCPSLP 766

## RESULT 9

US-10-295-027-922  
; Sequence 922, Application US/10295027  
; Publication No. US20030232350A1  
; GENERAL INFORMATION:  
; APPLICANT: Afar, Daniel  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Ginsberg, Wendy M.  
; APPLICANT: Gish, Kurt C.  
; APPLICANT: Glynn, Richard  
; APPLICANT: Hevezi, Peter A.  
; APPLICANT: Mack, David H.  
; APPLICANT: Murray, Richard  
; APPLICANT: Watson, Susan R.  
; APPLICANT: Eos Biotechnology, Inc.  
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
; FILE OF INVENTION: Methods of Screening for Modulators of Cancer  
; FILE REFERENCE: 018501-012500US  
; CURRENT APPLICATION NUMBER: US/10/295,027  
; CURRENT FILING DATE: 2002-11-13  
; PRIOR APPLICATION NUMBER: US 09/663,733  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: US 60/350,666  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/335,394  
; PRIOR FILING DATE: 2001-11-15

; PRIOR APPLICATION NUMBER: US 60/332,464  
; PRIOR FILING DATE: 2001-11-21  
; PRIOR APPLICATION NUMBER: US 60/334,393  
; PRIOR FILING DATE: 2001-11-29  
; PRIOR APPLICATION NUMBER: US 60/340,376  
; PRIOR FILING DATE: 2001-12-14  
; PRIOR APPLICATION NUMBER: US 60/347,211  
; PRIOR FILING DATE: 2002-01-08  
; PRIOR APPLICATION NUMBER: US 60/347,349  
; PRIOR FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 60/355,250  
; PRIOR FILING DATE: 2002-02-08  
; PRIOR APPLICATION NUMBER: US 60/356,714  
; PRIOR FILING DATE: 2002-02-13  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1386  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 922  
; LENGTH: 766  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-295-027-922  
Query Match 97.8%; Score 3933; DB 4; Length 766;  
Best Local Similarity 99.9%; Pred. No. 1.1e-305;  
Matches 727; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 13 SRKTYTLTDYLNKNTYRLKLYSLRWISDHEYLKQENNLVFNNAEYGNSSVFLNSTFDFP 72  
Db 39 SRKTYTLTDYLNKNTYRLKLYSLRWISDHEYLKQENNLVFNNAEYGNSSVFLNSTFDFP 98  
Qy 73 GHSINDYSIPDQFILLLEYNVVKQWRHSYASDYIDLKQRLTEERIPNNTQWVTWS 132  
Db 99 GHSINDYSIPDQFILLLEYNVVKQWRHSYASDYIDLKQRLTEERIPNNTQWVTWS 158  
Qy 133 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFSAYSALWSP 192  
Db 159 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFSAYSALWSP 218  
Qy 193 NGTEFLAYAQFNDTEVPLIEYSFYDESLOYPKTVRVPYKAGAVNPTVKFFVNTDLSLSS 252  
Db 219 NGTEFLAYAQFNDTEVPLIEYSFYDESLOYPKTVRVPYKAGAVNPTVKFFVNTDLSLSS 278  
Qy 253 VTNATSIQITAPASMLIGDHYLCVDTWATQERISLQWLRRIQNYSVMDICDYDESSGRWN 312  
Db 279 VTNATSIQITAPASMLIGDHYLCVDTWATQERISLQWLRRIQNYSVMDICDYDESSGRWN 338  
Qy 313 CLVARQHIEMSTTGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCCTFIT 372  
Db 339 CLVARQHIEMSTTGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCCTFIT 398  
Qy 373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCCLSELNPERCOYYS 432  
Db 399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCCLSELNPERCOYYS 458  
Qy 433 VSFSEAKYQYLRCSGPGPLLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPSSKKLDFII 492  
Db 459 VSFSEAKYQYLRCSGPGPLLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPSSKKLDFII 518  
Qy 493 LNETKFWYQMLPPHFDKSKYPLLLDVYAGPSCOKADTVFRLNWTYLASTENIIVASF 552  
Db 519 LNETKFWYQMLPPHFDKSKYPLLLDVYAGPSCOKADTVFRLNWTYLASTENIIVASF 578  
Qy 553 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGVVTSM 612  
Db 579 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGVVTSM 638  
Qy 613 VLGSYGVPKCGIAVAPVSRWEYDVSVYTERYMGLTPTPEDNLDHYNSTVMSRAENFKQV 672  
Db 639 VLGSYGVPKCGIAVAPVSRWEYDVSVYTERYMGLTPTPEDNLDHYNSTVMSRAENFKQV 698  
Qy 673 EYLLIHGTADDNVHFOQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHF 732



Db 699 EYLLHGHTADDNVHFQQAQISKALVDVGVDFQAMWYTTDEDHGIIASSTAHOHIYTHMSHP 758  
QY 733 IKQCFSLP 740  
Db 759 IKQCFSLP 766  
RESULT 10  
US-10-794-899-41  
; Sequence 41, Application US/10794899  
; Publication No. US20040146516A1  
; GENERAL INFORMATION:  
; APPLICANT: Utah Ventures  
; TITLE OF INVENTION: Lumen-Exposed Molecules and Methods Targeted Delivery  
; FILE REFERENCE: 27110-715  
; CURRENT APPLICATION NUMBER: US/10/794,899  
; CURRENT FILING DATE: 2004-03-05  
; NUMBER OF SEQ ID NOS: 121  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 41  
; LENGTH: 766  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-794-899-41  
Query Match 97.8%; Score 3933; DB 4; Length 766;  
Best Local Similarity 99.9%; Pred. No. 1.1e-305;  
Matches 727; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 13 SRKTYTLLDYLNKNTYRLKLSLRWISDHEYLKQENNLVFNAAEYGNSSVFLNSTPDEF 72  
Db 39 SRKTYTLLDYLNKNTYRLKLSLRWISDHEYLKQENNLVFNAAEYGNSSVFLNSTPDEF 98  
QY 73 GHSINDYSISPDGQFILLLEYVVKQWRHSYTSYDIYDLNKRQLITEERIPIINTQWVWS 132  
Db 99 GHSINDYSISPDGQFILLLEYVVKQWRHSYTSYDIYDLNKRQLITEERIPIINTQWVWS 158  
QY 133 PVGHKLAYVWNNDIYVVKIEPNLPSYRITWTCKEDIIYNGITDWVYEEVFSAYSALWWS 192  
Db 159 PVGHKLAYVWNNDIYVVKIEPNLPSYRITWTCKEDIIYNGITDWVYEEVFSAYSALWWS 218  
QY 193 NGTFLAYAQFNDTEVPLIEYSFYSDLSQYKPTVRVPYKAGANNPTKPFVWNTDSLSS 252  
Db 219 NGTFLAYAQFNDTEVPLIEYSFYSDLSQYKPTVRVPYKAGANNPTKPFVWNTDSLSS 278  
QY 253 VTNATSIQITAPASMLIGDHYLDCVWTATQBRISLOWLRRIONYSVMDICDYDESSGRWN 312  
Db 279 VTNATSIQITAPASMLIGDHYLDCVWTATQBRISLOWLRRIONYSVMDICDYDESSGRWN 338  
QY 313 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKXDCFTIT 372  
Db 339 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKXDCFTIT 398  
QY 373 KGTWEVIGIEALTSDYLYIISNEYKMGPGGNLYKIQLSDYTKVTCLSCELNPERCQYYS 432  
Db 399 KGTWEVIGIEALTSDYLYIISNEYKMGPGGNLYKIQLSDYTKVTCLSCELNPERCQYYS 458  
QY 433 VSPFSEKAYYQLRCSGPGPLPLYTLHSSVNDKGLVLEDNSALDKMLQNVQWPSKCLDFII 492  
Db 459 VSPFSEKAYYQLRCSGPGPLPLYTLHSSVNDKGLVLEDNSALDKMLQNVQWPSKCLDFII 518  
QY 493 LNETKFWYQMLPPHFDKSKYPLLLDVYAGPCSQKADTVFRLNWTATYLASTENIIVASF 552  
Db 519 LNETKFWYQMLPPHFDKSKYPLLLDVYAGPCSQKADTVFRLNWTATYLASTENIIVASF 578  
QY 553 DGRSGYQGDKIMHAINRLRGTFFVEDQIEAARFQKMGFVNDKRIAIWGSYGYVTSM 612  
Db 579 DGRSGYQGDKIMHAINRLRGTFFVEDQIEAARFQKMGFVNDKRIAIWGSYGYVTSM 638  
QY 613 VLGSYGKVFCKGIAVAPSRWEYDYSVYTERYMGLPEDNLDHYRNSTVMSRAENFKQV 672  
Db 639 VLGSYGKVFCKGIAVAPSRWEYDYSVYTERYMGLPEDNLDHYRNSTVMSRAENFKQV 698

QY 673 EYLLHGHTADDNVHFQQAQISKALVDVGVDFQAMWYTTDEDHGIIASSTAHOHIYTHMSHP 732  
Db 699 EYLLHGHTADDNVHFQQAQISKALVDVGVDFQAMWYTTDEDHGIIASSTAHOHIYTHMSHP 758  
QY 733 IKQCFSLP 740  
Db 759 IKQCFSLP 766  
RESULT 11  
US-11-041-674-6  
; Sequence 6, Application US/11041674  
; Publication No. US20050181468A1  
; GENERAL INFORMATION:  
; APPLICANT: Vanderbilt University  
; APPLICANT: Brown, Nancy J.  
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERT  
; TITLE OF INVENTION: INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA  
; FILE REFERENCE: Atty Docket No. 1242/48/2/2/2  
; CURRENT APPLICATION NUMBER: US/11/041,674  
; CURRENT FILING DATE: 2005-01-24  
; PRIOR APPLICATION NUMBER: 60/244,524  
; PRIOR FILING DATE: 2000-10-31  
; PRIOR APPLICATION NUMBER: 10/002,593  
; PRIOR FILING DATE: 2001-10-31  
; PRIOR APPLICATION NUMBER: 10/423,714  
; PRIOR FILING DATE: 2003-4-25  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 766  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-041-674-6  
Query Match 97.8%; Score 3933; DB 6; Length 766;  
Best Local Similarity 99.9%; Pred. No. 1.1e-305;  
Matches 727; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 13 SRKTYTLLDYLNKNTYRLKLSLRWISDHEYLKQENNLVFNAAEYGNSSVFLNSTPDEF 72  
Db 39 SRKTYTLLDYLNKNTYRLKLSLRWISDHEYLKQENNLVFNAAEYGNSSVFLNSTPDEF 98  
QY 73 GHSINDYSISPDGQFILLLEYVVKQWRHSYTSYDIYDLNKRQLITEERIPIINTQWVWS 132  
Db 99 GHSINDYSISPDGQFILLLEYVVKQWRHSYTSYDIYDLNKRQLITEERIPIINTQWVWS 158  
QY 133 PVGHKLAYVWNNDIYVVKIEPNLPSYRITWTCKEDIIYNGITDWVYEEVFSAYSALWWS 192  
Db 159 PVGHKLAYVWNNDIYVVKIEPNLPSYRITWTCKEDIIYNGITDWVYEEVFSAYSALWWS 218  
QY 193 NGTFLAYAQFNDTEVPLIEYSFYSDLSQYKPTVRVPYKAGANNPTKPFVWNTDSLSS 252  
Db 219 NGTFLAYAQFNDTEVPLIEYSFYSDLSQYKPTVRVPYKAGANNPTKPFVWNTDSLSS 278  
QY 253 VTNATSIQITAPASMLIGDHYLDCVWTATQBRISLOWLRRIONYSVMDICDYDESSGRWN 312  
Db 279 VTNATSIQITAPASMLIGDHYLDCVWTATQBRISLOWLRRIONYSVMDICDYDESSGRWN 338  
QY 313 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKXDCFTIT 372  
Db 339 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKXDCFTIT 398  
QY 373 KGTWEVIGIEALTSDYLYIISNEYKMGPGGNLYKIQLSDYTKVTCLSCELNPERCQYYS 432  
Db 399 KGTWEVIGIEALTSDYLYIISNEYKMGPGGNLYKIQLSDYTKVTCLSCELNPERCQYYS 458  
QY 433 VSPFSEKAYYQLRCSGPGPLPLYTLHSSVNDKGLVLEDNSALDKMLQNVQWPSKCLDFII 492  
Db 459 VSPFSEKAYYQLRCSGPGPLPLYTLHSSVNDKGLVLEDNSALDKMLQNVQWPSKCLDFII 518  
QY 493 LNETKFWYQMLPPHFDKSKYPLLLDVYAGPCSQKADTVFRLNWTATYLASTENIIVASF 552

Db 519 LNETHFWYQMLPPHFDKSKYPLLLDVYAGPCKQADTVFRLNWNATYLASTENIIVASF 578  
QY 553 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGVVTSM 612  
Db 579 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGVVTSM 638  
QY 613 VLGSGGVFKGCIAPVSRWEYDSVYTERYMGFLPTPEDNLDHYRNSVTWMSRAENFKQV 672  
Db 639 VLGSGGVFKGCIAPVSRWEYDSVYTERYMGFLPTPEDNLDHYRNSVTWMSRAENFKQV 698  
QY 673 EYLLHGTADDNVHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHF 732  
Db 699 EYLLHGTADDNVHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHF 758  
QY 733 IKQCFSLP 740  
Db 759 IKQCFSLP 766

## RESULT 12

US-10-476-264-147  
; Sequence 147, Application US/10476264  
; Publication No. US20050123910A1  
; GENERAL INFORMATION:  
; APPLICANT: Cookeon, William Osmond Charles Michael  
; APPLICANT: Moffat, Miriam Fleur  
; APPLICANT: Allen, Maxine  
; APPLICANT: Lench, Nick  
; TITLE OF INVENTION: Enzyme and SNP marker for disease  
; FILE REFERENCE: 16721-002US1  
; CURRENT APPLICATION NUMBER: US/10/476,264  
; CURRENT FILING DATE: 2003-10-24  
; PRIOR APPLICATION NUMBER: PCT/GB02/01887  
; PRIOR FILING DATE: 2002-04-24  
; PRIOR APPLICATION NUMBER: GB0110044.5  
; PRIOR FILING DATE: 2001-04-24  
; PRIOR APPLICATION NUMBER: GB0110046.0  
; PRIOR FILING DATE: 2001-04-24  
; PRIOR APPLICATION NUMBER: GB0124594.3  
; PRIOR FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: GB0124575.2  
; PRIOR FILING DATE: 2001-10-12  
; NUMBER OF SEQ ID NOS: 421  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 147  
; LENGTH: 766  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-476-264-147

Query Match 97.7%; Score 3929; DB 5; Length 766;  
Best Local Similarity 99.7%; Pred. No. 2.3e-305;  
Matches 726; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 13 SRKTYTLTDYLNKTYRLKLSLRWISDHEYLKQENNILVFNAEYGNSSVFLNENSTFDEF 72  
Db 39 SRKTYTLTDYLNKTYRLKLSLRWISDHEYLKQENNILVFNAEYGNSSVFLNENSTFDEF 98  
QY 73 GHSINDYSISPDQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERI PNTQWVTWS 132  
Db 99 GHSINDYSISPDQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERI PNTQWVTWS 158  
QY 133 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWMVEEVEFSAYSALWWSF 192  
Db 159 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWMVEEVEFSAYSALWWSF 218  
QY 193 NGTFLEYAQFNDTEVPLIEYSFYSDESLOYPKTVRVPYKAGAVNPTVKFFVNTDLSLS 252  
Db 219 NGTFLEYAQFNDTEVPLIEYSFYSDESLOYPKTVRVPYKAGAVNPTVKFFVNTDLSLS 278  
QY 253 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLQWLRRIQNYSYMDICDYDESSGRWN 312  
Db 279 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLQWLRRIQNYSYMDICDYDESSGRWN 338

QY 313 CLVARQHLEMTSTGWGRFRPSEPHFTLDGNSFYKIISNEEGYRHICYFOIDKDCCTFIT 372  
Db 339 CLVARQHLEMTSTGWGRFRPSEPHFTLDGNSFYKIISNEEGYRHICYFOIDKDCCTFIT 398  
QY 373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS 432  
Db 399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS 458  
QY 433 VSFSKEAKYQLRCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFFII 492  
Db 459 VSFSKEAKYQLRCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFFII 518  
QY 493 LNETHFWYQMLPPHFDKSKYPLLLDVYAGPCKQADTVFRLNWNATYLASTENIIVASF 552  
Db 519 LNETHFWYQMLPPHFDKSKYPLLLDVYAGPCKQADTVFRLNWNATYLASTENIIVASF 578  
QY 553 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGVVTSM 612  
Db 579 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGVVTSM 638  
QY 613 VLGSGGVFKGCIAPVSRWEYDSVYTERYMGFLPTPEDNLDHYRNSVTWMSRAENFKQV 672  
Db 639 VLGSGGVFKGCIAPVSRWEYDSVYTERYMGFLPTPEDNLDHYRNSVTWMSRAENFKQV 698  
QY 673 EYLLHGTADDNVHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHF 732  
Db 699 EYLLHGTADDNVHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHF 758  
QY 733 IKQCFSLP 740  
Db 759 IKQCFSLP 766

## RESULT 13

US-10-770-712-1  
; Sequence 1, Application US/10770712  
; Publication No. US20050170333A1  
; GENERAL INFORMATION:  
; APPLICANT: Vojdani, Aristo  
; TITLE OF INVENTION: IDENTIFICATION OF ETIOLOGY OF AUTISM  
; FILE REFERENCE: IMSC12.008A  
; CURRENT APPLICATION NUMBER: US/10/770,712  
; CURRENT FILING DATE: 2004-02-03  
; NUMBER OF SEQ ID NOS: 133  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 766  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-770-712-1

Query Match 97.7%; Score 3929; DB 5; Length 766;  
Best Local Similarity 99.7%; Pred. No. 2.3e-305;  
Matches 726; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 13 SRKTYTLTDYLNKTYRLKLSLRWISDHEYLKQENNILVFNAEYGNSSVFLNENSTFDEF 72  
Db 39 SRKTYTLTDYLNKTYRLKLSLRWISDHEYLKQENNILVFNAEYGNSSVFLNENSTFDEF 98  
QY 73 GHSINDYSISPDQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERI PNTQWVTWS 132  
Db 99 GHSINDYSISPDQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERI PNTQWVTWS 158  
QY 133 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWMVEEVEFSAYSALWWSF 192  
Db 159 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWMVEEVEFSAYSALWWSF 218  
QY 193 NGTFLEYAQFNDTEVPLIEYSFYSDESLOYPKTVRVPYKAGAVNPTVKFFVNTDLSLS 252  
Db 219 NGTFLEYAQFNDTEVPLIEYSFYSDESLOYPKTVRVPYKAGAVNPTVKFFVNTDLSLS 278  
QY 253 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLQWLRRIQNYSYMDICDYDESSGRWN 312

Db 279 VTNATSIQITAPASMLIGDHYLCVDTWATQBRISLQWLRRIQNYSMVMDICDYDESSGRWN 338  
Qy 313 CLVARQHIEMSTTGVGGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKCKCTIT 372  
Db 339 CLVARQHIEMSTTGVGGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKCKCTIT 398  
Qy 373 KGTWEVIGIEALTSDYLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYYS 432  
Db 399 KGTWEVIGIEALTSDYLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYYS 458  
Qy 433 VSFSEKAYYQLRCSGPGPLPYTLHSSVNDKGLRVLEDSALDKMLQNVQMPKSKLDPII 492  
Db 459 VSFSEKAYYQLRCSGPGPLPYTLHSSVNDKGLRVLEDSALDKMLQNVQMPKSKLDPII 518  
Qy 493 LNETFWYQMLIPPHFDKSKKYPLLDVYAGPCSKADIVPRLNWTATYLASTENIIVASF 552  
Db 519 LNETFWYQMLIPPHFDKSKKYPLLDVYAGPCSKADIVPRLNWTATYLASTENIIVASF 578  
Qy 553 DGRSGYQGDKIMHAINRRLGTFFVEDQIEAARQFSKMGFVDNKRKIAIWGHSYGGYVTS 612  
Db 579 DGRSGYQGDKIMHAINRRLGTFFVEDQIEAARQFSKMGFVDNKRKIAIWGHSYGGYVTS 638  
Qy 613 VLGSYGVPKCGIAVAPVSRWEYDVSVTERYMGGLPTPEDNLDHYRNSTVMSRAENPQV 672  
Db 639 VLGSYGVPKCGIAVAPVSRWEYDVSVTERYMGGLPTPEDNLDHYRNSTVMSRAENPQV 698  
Qy 673 EYLLIHGTADDNVHFQQAQISKALVDVGVDFQAMWYTDHGIASSTAHQHIYTHMSHF 732  
Db 699 EYLLIHGTADDNVHFQQAQISKALVDVGVDFQAMWYTDHGIASSTAHQHIYTHMSHF 758  
Qy 733 IKQCFSLP 740  
Db 759 IKQCFSLP 766

## RESULT 14

US-10-770-712-2  
; Sequence 2, Application US/10770712  
; Publication No. US20050170333A1  
; GENERAL INFORMATION:  
; APPLICANT: Vojdani, Aristo  
; TITLE OF INVENTION: IDENTIFICATION OF ETIOLOGY OF AUTISM  
; FILE REFERENCE: IMSC12.008A  
; CURRENT APPLICATION NUMBER: US/10770,712  
; CURRENT FILING DATE: 2004-02-03  
; NUMBER OF SEQ ID NOS: 133  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 767  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-10-770-712-2

Query Match 84.7%; Score 3406.5; DB 5; Length 767;  
Best Local Similarity 84.7%; Pred. No. 1.6e-263;  
Matches 618; Conservative 52; Mismatches 57; Indels 3; Gaps 1;

Qy 13 SRKTYTLTKNTYRLKLSLRWISDHELYLKQENNILVFNAYGNSVPLENSTDEF 72  
Db 37 SARTYTLADYLNKTFRVKYSLSRWVSDSEYLYKQENNILVFNAYGNSVPLENSTFEIF 96  
Qy 73 GHSINDYSISPDGQFILLLEYNVQWRHSYTSYDIYDLNKRQLITTEERIPNNTQWITS 132  
Db 97 GDSISDSVSPDRFLVLEYNVQWRHSYTSYDIYDLNKRQLITTEERIPNNTQWITS 156  
Qy 133 PVGHKLAYVWNNNDIYVVKIEPLPSYRITWTGKEDIYNGITDMVYBEVSAYSALWSP 192  
Db 157 QEGHKLAYVWNNNDIYVVKIEPLPSYRITWTGKEDIYNGITDMVYBEVSAYSALWSP 216  
Qy 193 NGTFLAYAQFNDTVPVLEIYSFYDSLSLQYPTKVVPYKAGAVNPTKFPVNTDSLS 252  
Db 217 NGTFLAYAQFNDTVPVLEIYSFYDSLSLQYPTKVVPYKAGAVNPTKFPVNTDSLS 276

Qy 253 VTNATSIQITAPASMLIGDHYLCVDTWATQBRISLQWLRRIQNYSMVMDICDYDESSGRWN 312  
Db 277 TTTTTPMQITAPASVTTGDHYLCVAVWVSEDRIISQWLRRIQNYSMVMDICDYDESSGRWN 336  
Qy 313 CLVARQHIEMSTTGVGGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKCKD---CT 369  
Db 337 CPTTREHIEATSGWGRFRPSEPHFTSDGSSFYKIVSDKGYKHICQFQDKRKEPVCT 396  
Qy 370 FITKGTWEVIGIEALTSDYLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQ 429  
Db 397 FITKGTWEVIGIEALTSDYLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQ 456  
Qy 430 YYSVSFSKAYYQLRCSGPGPLPYTLHSSVNDKGLRVLEDSALDKMLQNVQMPKSKLD 489  
Db 457 YYSVSFSKAYYQLRCSGPGPLPYTLHSSVNDKGLRVLEDSALDKMLQNVQMPKSKLD 516  
Qy 490 FIILNETFWYQMLIPPHFDKSKKYPLLDVYAGPCSKADIVPRLNWTATYLASTENIIV 549  
Db 517 FIVLNETFWYQMLIPPHFDKSKKYPLLDVYAGPCSKADIVPRLNWTATYLASTENIIV 576  
Qy 550 ASPDGRSGYQGDKIMHAINRRLGTFFVEDQIEAARQFSKMGFVDNKRKIAIWGHSYGGV 609  
Db 577 ASPDGRSGYQGDKIMHAINRRLGTFFVEDQIEAARQFSKMGFVDNKRKIAIWGHSYGGV 636  
Qy 610 TSMVLGSYGVPKCGIAVAPVSRWEYDVSVTERYMGGLPTPEDNLDHYRNSTVMSRAENP 669  
Db 637 TSMVLGSYGVPKCGIAVAPVSRWEYDVSVTERYMGGLPTPEDNLDHYRNSTVMSRAENP 696  
Qy 670 KQVEYLLIHGTADDNVHFQQAQISKALVDVGVDFQAMWYTDHGIASSTAHQHIYTHM 729  
Db 697 KQVEYLLIHGTADDNVHFQQAQISKALVDVGVDFQAMWYTDHGIASSTAHQHIYTHM 756  
Qy 730 SHFIKQCFSL 739  
Db 757 SHFIKQCFSL 766

## RESULT 15

US-10-165-603-4  
; Sequence 4, Application US/10165603  
; Publication No. US20030021792A1  
; GENERAL INFORMATION:  
; APPLICANT: Roben, Paul W.  
; TITLE OF INVENTION: TISSUE-SPECIFIC ENDOTHELIAL MEMBRANE  
; TITLE OF INVENTION: PROTEINS  
; FILE REFERENCE: TPTECH.001A  
; CURRENT APPLICATION NUMBER: US/10/165,603  
; CURRENT FILING DATE: 2002-06-07  
; PRIOR APPLICATION NUMBER: 60/297,021  
; PRIOR FILING DATE: 2001-06-08  
; PRIOR APPLICATION NUMBER: 60/305,117  
; PRIOR FILING DATE: 2001-07-12  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 767  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-10-165-603-4

Query Match 84.5%; Score 3395.5; DB 4; Length 767;  
Best Local Similarity 84.4%; Pred. No. 1.2e-262;  
Matches 616; Conservative 53; Mismatches 58; Indels 3; Gaps 1;

Qy 13 SRKTYTLTKNTYRLKLSLRWISDHELYLKQENNILVFNAYGNSVPLENSTDEF 72  
Db 37 SARTYTLADYLNKTFRVKYSLSRWVSDSEYLYKQENNILVFNAYGNSVPLENSTFEIF 96  
Qy 73 GHSINDYSISPDGQFILLLEYNVQWRHSYTSYDIYDLNKRQLITTEERIPNNTQWITS 132  
Db 97 GDSISDSVSPDRFLVLEYNVQWRHSYTSYDIYDLNKRQLITTEERIPNNTQWITS 156

```

QY 133 PVCHKLAYVWNDIYVYKIEPNLPSYRIITWTGKEDIYNGITDQVYEEVEFSAYSALWSP 192
Db 157 QEGHKLAYVWKNDIYVYKIEPHLPSHRITSTGENVIENGINDQVYEEVEFGAYSALWSP 216
QY 193 NGTFLAYAQFNDETEVPLIEYSFYSDSLQYPKTVRPYPKAGAVNPTVKFFVNTDSLSS 252
Db 217 NGTFLAYAQFNDETEVPLIEYSFYSDSLQYPKTVRPYPKAGAVNPTVKFFVNTDSLSS 276
QY 253 VTNATSIQITAPASMLIGHYLCVDTWATQERISQWLRRIQNYSDICDYDESSGRWN 312
Db 277 TTTTIPMQITAPASVTTGDHYLCDVAWSEDRISQWLRRIQNYSWAICDYDKTTLVWN 336
QY 313 CLVAQCHIEMSITGWRFRPPEPHFTLDGNSFYKIIISNEEGYRHICYFOIDKKD---CT 369
Db 337 CPTTQEHIEISATGWCGRFRPAPPHFTSDGSSFYKIVSDKGYKHICQFQDKRKEQVCT 396
QY 370 FITKGTWEVIGIEALTSYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQ 429
Db 397 FITKGAWVISIEALTSYLYISNEYKMGPGGRNLYKIQLTDHTNKKCLSCDLNPERCQ 456
QY 430 YYSVSFSKAKYYQLRCSGGLPLTYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLD 489
Db 457 YYSVLSKEAKYYGLGCRGGLPLTYTLHRSTDQKELRVLEDNSALDKMLQDVQMPSSKLD 516
QY 490 FIILNETKFWYQMLPPHEDKSKYPLLLDVYAGPCQKADTVPRLNWATYLASTENIIV 549
Db 517 FIVLNETRFYQMLPPHFDKSKYPLLLDVYAGPCQKADAAFLRNWATYLASTENIIV 576
QY 550 ASFDGRGSGYQGDKIMHAINRRLGTPEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGYV 609
Db 577 ASFDGRGSGYQGDKIMHAINRRLGTLEVEDQIEAARQFLKMGFVDSKRVAIWGSYGGYV 636
QY 610 TSMVLGSGGVPKCGIAVAPVSRWEYYSYTERYMGGLPTPEDNLDHYRNSTVMSRAENF 669
Db 637 TSMVLGSGGVPKCGIAVAPVSRWEYYSYTERYMGGLPTPEDNLDHYRNSTVMSRAENF 696
QY 670 KQVEYLLIHGTADDNVHFQCSAQISKALVDVGVDFQAMWYTTDEDHGIASSTAHOHIYTHM 729
Db 697 KQVEYLLIHGTADDNVHFQCSAQISKALVDAGVDFQAMWYTTDEDHGIASSTAHOHIYSHM 756
QY 730 SHPIKOCFSL 739
Db 757 SHFLQOCFSL 766

```

Search completed: February 15, 2006, 13:29:19  
Job time : 171 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2006, 13:14:53 ; Search time 48 Seconds  
(without alignments)  
1274.583 Million cell updates/sec

Title: US-10-659-055-3  
Perfect score: 4020  
Sequence: 1 ADPGSHHHHRRKTYTLT.....AQHIYTHMSHFQKCFSLP 740

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA.\*  
1: /cgm2\_6/ptodata/1/iaa/5 COMB.pap.\*  
2: /cgm2\_6/ptodata/1/iaa/6 COMB.pap.\*  
3: /cgm2\_6/ptodata/1/iaa/H COMB.pap.\*  
4: /cgm2\_6/ptodata/1/iaa/PCTUS COMB.pap.\*  
5: /cgm2\_6/ptodata/1/iaa/RE COMB.pap.\*  
6: /cgm2\_6/ptodata/1/iaa/backfiles.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3939	98.0	755	4	PCT-US93-07923-3
2	3939	98.0	759	4	PCT-US93-07923-2
3	3939	98.0	766	1	US-08-230-491A-3
4	3939	98.0	766	1	US-08-619-280A-3
5	3939	98.0	766	1	US-08-940-391-3
6	3939	98.0	766	2	US-09-794-236-1
7	3939	98.0	766	2	US-09-265-606-3
8	3939	98.0	775	2	US-09-949-016-10450
9	3933	97.8	766	2	US-10-002-593-6
10	3933	97.8	766	2	US-09-949-016-6146
11	3933	97.8	766	2	US-10-423-714-6
12	3929	97.7	766	2	US-09-518-550-27
13	3010	74.9	593	4	PCT-US93-07923-11
14	2168	53.9	760	1	US-08-230-491A-2
15	2168	53.9	760	1	US-08-619-280A-2
16	2168	53.9	760	1	US-08-940-391-2
17	2168	53.9	760	2	US-09-265-606-2
18	2158.5	53.7	761	2	US-09-518-550-26
19	1259.5	31.3	547	2	US-09-949-016-8330
20	1223	30.4	796	2	US-09-976-674-5
21	1198	29.8	743	2	US-10-363-937-4
22	1196	29.8	706	2	US-09-976-674-41
23	1152.5	28.7	691	2	US-09-976-674-43
24	987	24.6	771	2	US-09-079-592-2
25	963.5	24.0	771	2	US-09-462-284-2
26	914	22.7	818	2	US-09-462-845-3
27	914	22.7	818	2	US-10-402-312-3

28	914	22.7	818	2	US-10-401-437-3	Sequence 3, Appli
29	914	22.7	818	2	US-10-402-067-3	Sequence 3, Appli
30	914	22.7	818	2	US-10-401-436-3	Sequence 3, Appli
31	809	20.1	676	2	US-09-518-550-42	Sequence 42, Appl
32	809	20.1	723	2	US-09-518-550-29	Sequence 29, Appl
33	795.5	19.9	710	2	US-09-518-550-28	Sequence 28, Appl
34	733	18.2	931	2	US-09-079-592-11	Sequence 11, Appl
35	591	14.7	323	2	US-09-270-767-45296	Sequence 45296, A
36	584.5	14.5	494	2	US-09-248-796A-19472	Sequence 19472, A
37	528.5	13.1	882	2	US-09-976-674-1	Sequence 1, Appli
38	528.5	13.1	882	2	US-10-070-464-1	Sequence 1, Appli
39	500	12.4	879	2	US-09-976-674-33	Sequence 33, Appl
40	500	12.4	879	2	US-09-976-674-35	Sequence 35, Appl
41	497	12.4	732	2	US-09-518-550-30	Sequence 30, Appl
42	493.5	12.3	526	2	US-09-248-796A-19924	Sequence 19924, A
43	493.5	12.3	863	2	US-09-976-674-3	Sequence 3, Appli
44	493.5	12.3	892	2	US-09-976-674-23	Sequence 23, Appl
45	493.5	12.3	892	2	US-09-976-674-27	Sequence 27, Appl

ALIGNMENTS

RESULT 1  
PCT-US93-07923-3  
; Sequence 3, Application PC/TUS9307923  
; GENERAL INFORMATION:  
; APPLICANT: Morimoto, Chikao  
; APPLICANT: Tanaka, Toshiaki  
; TITLE OF INVENTION: HUMAN CD26 AND METHODS FOR USE  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 502 or 55SX  
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
; SOFTWARE: WordPerfect (Version 5.0)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/07923  
; FILING DATE: 19930819  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/934,162  
; FILING DATE: 21-AUG-1992  
; APPLICATION NUMBER: 07/832,211  
; FILING DATE: 06-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fraser, Janis K.  
; REGISTRATION NUMBER: 34,819  
; REFERENCE/DOCKET NUMBER: 00530/055002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 755  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
PCT-US93-07923-3

Query Match 98.0%; Score 3939; DB 4; Length 755;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 13 SRKTYTLDTYLNKTYVRLKLYSLRWISDHEYLKQENNIILVFNABEYGNSSVFLENSTFDEF 72
Db 28 SRKTYTLDTYLNKTYVRLKLYSLRWISDHEYLKQENNIILVFNABEYGNSSVFLENSTFDEF 87
QY 73 GHSINDYSISPDQGFILLEYNYVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQVWTWS 132
Db 88 GHSINDYSISPDQGFILLEYNYVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQVWTWS 147
QY 133 PVGHKLAYVNNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFSAYSALWWSF 192
Db 148 PVGHKLAYVNNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFSAYSALWWSF 207
QY 193 NGTFELAYAQFNDTEVPLEIYSFYSDESLOYPKTVRPVYPKAGAVNPTVKFFVNTDLSLS 252
Db 208 NGTFELAYAQFNDTEVPLEIYSFYSDESLOYPKTVRPVYPKAGAVNPTVKFFVNTDLSLS 267
QY 253 VTNATSIQITAPASMLIGDHYLDCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWN 312
Db 268 VTNATSIQITAPASMLIGDHYLDCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWN 327
QY 313 CLVARQHIEEMSTTGVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKCKCTFIT 372
Db 328 CLVARQHIEEMSTTGVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKCKCTFIT 387
QY 373 KGTWEVIGIEALTSYLYIISNEYKMGPGRNLYKIQLSDYTKVTCLSCELNPERCOYYS 432
Db 388 KGTWEVIGIEALTSYLYIISNEYKMGPGRNLYKIQLSDYTKVTCLSCELNPERCOYYS 447
QY 433 VFSFKEAKYQLRCSGPGPLVYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDLFI 492
Db 448 VFSFKEAKYQLRCSGPGPLVYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDLFI 507
QY 493 LNETKFWYQMLPPHFDKSKYKPLLDVYAGPSCQKADTVFRLNWTATYLASTENIIVASF 552
Db 508 LNETKFWYQMLPPHFDKSKYKPLLDVYAGPSCQKADTVFRLNWTATYLASTENIIVASF 567
QY 553 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYVTSM 612
Db 568 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYVTSM 627
QY 613 VLGSQGVFKCGIAPVSRWEYDSYVYTERYNGLTPEDNLDHYNHNSTVMSRAENFKQV 672
Db 628 VLGSQGVFKCGIAPVSRWEYDSYVYTERYNGLTPEDNLDHYNHNSTVMSRAENFKQV 687
QY 673 EYLLIHGTADDNVHFQOQAISKALVDVGVDFQAMWYTTDEDHGIASSTAHOIYTHMSHF 732
Db 688 EYLLIHGTADDNVHFQOQAISKALVDVGVDFQAMWYTTDEDHGIASSTAHOIYTHMSHF 747
QY 733 IKQCFSLP 740
Db 748 IKQCFSLP 755

```

## RESULT 2

```

PCT-US93-07923-2
; Sequence 2, Application PC/TUS9307923
; GENERAL INFORMATION:
; APPLICANT: Morimoto, Chikao
; APPLICANT: Schlosgman, Stuart F.
; APPLICANT: Tanaka, Toshiaki
; TITLE OF INVENTION: HUMAN CD26 AND METHODS FOR USE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)

```

```

; SOFTWARE: WordPerfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07923
; FILING DATE: 19930819
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934,162
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: 07/832,211
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 759
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; PCT-US93-07923-2

Query Match 98.0%; Score 3939; DB 4; Length 759;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 SRKTYTLDTYLNKTYVRLKLYSLRWISDHEYLKQENNIILVFNABEYGNSSVFLENSTFDEF 72
Db 32 SRKTYTLDTYLNKTYVRLKLYSLRWISDHEYLKQENNIILVFNABEYGNSSVFLENSTFDEF 91
QY 73 GHSINDYSISPDQGFILLEYNYVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQVWTWS 132
Db 92 GHSINDYSISPDQGFILLEYNYVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQVWTWS 151
QY 133 PVGHKLAYVNNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFSAYSALWWSF 192
Db 152 PVGHKLAYVNNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFSAYSALWWSF 211
QY 193 NGTFELAYAQFNDTEVPLEIYSFYSDESLOYPKTVRPVYPKAGAVNPTVKFFVNTDLSLS 252
Db 212 NGTFELAYAQFNDTEVPLEIYSFYSDESLOYPKTVRPVYPKAGAVNPTVKFFVNTDLSLS 271
QY 253 VTNATSIQITAPASMLIGDHYLDCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWN 312
Db 272 VTNATSIQITAPASMLIGDHYLDCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWN 331
QY 313 CLVARQHIEEMSTTGVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKCKCTFIT 372
Db 332 CLVARQHIEEMSTTGVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKCKCTFIT 391
QY 373 KGTWEVIGIEALTSYLYIISNEYKMGPGRNLYKIQLSDYTKVTCLSCELNPERCOYYS 432
Db 392 KGTWEVIGIEALTSYLYIISNEYKMGPGRNLYKIQLSDYTKVTCLSCELNPERCOYYS 451
QY 433 VFSFKEAKYQLRCSGPGPLVYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDLFI 492
Db 452 VFSFKEAKYQLRCSGPGPLVYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDLFI 511
QY 493 LNETKFWYQMLPPHFDKSKYKPLLDVYAGPSCQKADTVFRLNWTATYLASTENIIVASF 552
Db 512 LNETKFWYQMLPPHFDKSKYKPLLDVYAGPSCQKADTVFRLNWTATYLASTENIIVASF 571
QY 553 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYVTSM 612
Db 572 DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYVTSM 631
QY 613 VLGSQGVFKCGIAPVSRWEYDSYVYTERYNGLTPEDNLDHYNHNSTVMSRAENFKQV 672
Db 632 VLGSQGVFKCGIAPVSRWEYDSYVYTERYNGLTPEDNLDHYNHNSTVMSRAENFKQV 691

```





```

; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-940-391-3

Query Match      98.0%; Score 3939; DB 1; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 SRKTYTLDTYLNKTYRLKLYSLRWISDHEYLKQENNLVFNABYGNSSVFLENSTFDF 72
Db 39 SRKTYTLDTYLNKTYRLKLYSLRWISDHEYLKQENNLVFNABYGNSSVFLENSTFDF 98
QY 73 GHSINDYSIPGQFILLKYVNVKQWHSYTSASYDIYDLNKRQLITEERIPNNTQVTV 132
Db 99 GHSINDYSIPGQFILLKYVNVKQWHSYTSASYDIYDLNKRQLITEERIPNNTQVTV 158
QY 133 PVGHKLAVYNNNDIYVKIEPNLPSYRIWTGKEDIYNGITDWVYEEVFSAYSALWSP 192
Db 159 PVGHKLAVYNNNDIYVKIEPNLPSYRIWTGKEDIYNGITDWVYEEVFSAYSALWSP 218
QY 193 NGTFLAYAQFNDETEVPLIEYSFYSDLSQYPKTVRPVPKAGAVNPTVKFFVNTDSL 252
Db 219 NGTFLAYAQFNDETEVPLIEYSFYSDLSQYPKTVRPVPKAGAVNPTVKFFVNTDSL 278
QY 253 VTNATSIQITAPASMLIGDHYLCVDTWATQERISIQWLRRIONYSVMDICDYDESS 312
Db 279 VTNATSIQITAPASMLIGDHYLCVDTWATQERISIQWLRRIONYSVMDICDYDESS 338
QY 313 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCFT 372
Db 339 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCFT 398
QY 373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLKYQLSDYTKVTCLSCELNPERCQ 432
Db 399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLKYQLSDYTKVTCLSCELNPERCQ 458
QY 433 VSFSKEAKYQLRCSGPGILPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMP 492
Db 459 VSFSKEAKYQLRCSGPGILPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMP 518
QY 493 LNETKFWYQMLPPHFDKSKYPLLDVYAGCSQKADTVFRLNWTATYLASTENIIVASF 552
Db 519 LNETKFWYQMLPPHFDKSKYPLLDVYAGCSQKADTVFRLNWTATYLASTENIIVASF 578
QY 553 DGRGSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVGNKRIAIWGSYGGY 612
Db 579 DGRGSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVGNKRIAIWGSYGGY 638
QY 613 VLGSGSGVFKCIGIAPVSRWEYDVSVYTERYMGFLPTPEDNLHDYRNSVTMSRA 672
Db 639 VLGSGSGVFKCIGIAPVSRWEYDVSVYTERYMGFLPTPEDNLHDYRNSVTMSRA 698
QY 673 EYLLHGTADDNVHFOQSAISKALVDGVDFQAMWYTDDEHGIASSTAHOIYTHMSHF 732
Db 699 EYLLHGTADDNVHFOQSAISKALVDGVDFQAMWYTDDEHGIASSTAHOIYTHMSHF 758
QY 733 IKQCFSLP 740
Db 759 IKQCFSLP 766
```

RESULT 5  
US-08-940-391-3  
Sequence 3, Application US/08940391  
Patent No. 5965373  
GENERAL INFORMATION:  
APPLICANT: Zimmermann, Rainer; Park, John E.;  
Rettig, Wolfgang; Old, Lloyd J.  
TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION  
PROTEIN ALPHA, AND USES THEREOF  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:

```

; ADDRESS: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/940,391
; FILING DATE: 01-OCT-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/619,280
; FILING DATE: 18-MARCH-1996
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5965373man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-940-391-3
```

Query Match 98.0%; Score 3939; DB 1; Length 766;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 SRKTYTLDTYLNKTYRLKLYSLRWISDHEYLKQENNLVFNABYGNSSVFLENSTFDF 72  
Db 39 SRKTYTLDTYLNKTYRLKLYSLRWISDHEYLKQENNLVFNABYGNSSVFLENSTFDF 98  
QY 73 GHSINDYSIPGQFILLKYVNVKQWHSYTSASYDIYDLNKRQLITEERIPNNTQVTV 132  
Db 99 GHSINDYSIPGQFILLKYVNVKQWHSYTSASYDIYDLNKRQLITEERIPNNTQVTV 158  
QY 133 PVGHKLAVYNNNDIYVKIEPNLPSYRIWTGKEDIYNGITDWVYEEVFSAYSALWSP 192  
Db 159 PVGHKLAVYNNNDIYVKIEPNLPSYRIWTGKEDIYNGITDWVYEEVFSAYSALWSP 218  
QY 193 NGTFLAYAQFNDETEVPLIEYSFYSDLSQYPKTVRPVPKAGAVNPTVKFFVNTDSL 252  
Db 219 NGTFLAYAQFNDETEVPLIEYSFYSDLSQYPKTVRPVPKAGAVNPTVKFFVNTDSL 278  
QY 253 VTNATSIQITAPASMLIGDHYLCVDTWATQERISIQWLRRIONYSVMDICDYDESS 312  
Db 279 VTNATSIQITAPASMLIGDHYLCVDTWATQERISIQWLRRIONYSVMDICDYDESS 338  
QY 313 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCFT 372  
Db 339 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCFT 398  
QY 373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLKYQLSDYTKVTCLSCELNPERCQ 432  
Db 399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLKYQLSDYTKVTCLSCELNPERCQ 458  
QY 433 VSFSKEAKYQLRCSGPGILPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMP 492  
Db 459 VSFSKEAKYQLRCSGPGILPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMP 518  
QY 493 LNETKFWYQMLPPHFDKSKYPLLDVYAGCSQKADTVFRLNWTATYLASTENIIVASF 552  
Db 519 LNETKFWYQMLPPHFDKSKYPLLDVYAGCSQKADTVFRLNWTATYLASTENIIVASF 578

QY 553 DGRSGYQDQKIMHAINRRLGTFFVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVVTSM 612  
DB 579 DGRSGYQDQKIMHAINRRLGTFFVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVVTSM 638  
QY 613 VLGSYGVPKCGIAVAPVSRWEYDVSVTERYMGLTPTEDNLDHYRNSTVMSRAENFKOV 672  
DB 639 VLGSYGVPKCGIAVAPVSRWEYDVSVTERYMGLTPTEDNLDHYRNSTVMSRAENFKOV 698  
QY 673 EYLLHGTADDNVHFQQAQISKALVDGVDFQAMWYTDDEHGIASSTAHQHIYTHMSHF 732  
DB 699 EYLLHGTADDNVHFQQAQISKALVDGVDFQAMWYTDDEHGIASSTAHQHIYTHMSHF 758  
QY 733 IKQCFSLP 740  
DB 759 IKQCFSLP 766  
RESULT 6  
US-09-794-236-1  
; Sequence 1, Application US/09794236  
; Patent No. 6337069  
; GENERAL INFORMATION:  
; APPLICANT: Grouzmann, Eric  
; APPLICANT: Lacroix, Jean-Silvain  
; APPLICANT: Monod, Michel  
; TITLE OF INVENTION: Method of Treating Rhinitis and Sinusitis  
; FILE REFERENCE: 81985/276823  
; CURRENT APPLICATION NUMBER: US/09/794,236  
; CURRENT FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 766  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-794-236-1

Query Match 98.0%; Score 3939; DB 2; Length 766;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 13 SRKTYTLDYLNKTYRLKLSLRWISDHEYLKQENNILVFNAYGNSVFLNSTDEF 72  
DB 39 SRKTYTLDYLNKTYRLKLSLRWISDHEYLKQENNILVFNAYGNSVFLNSTDEF 98  
QY 73 GHSINDYISPDGQFILLVYVQWRHSYTSYDIYDLNKRQLITEERIPNTQVWTS 132  
DB 99 GHSINDYISPDGQFILLVYVQWRHSYTSYDIYDLNKRQLITEERIPNTQVWTS 158  
QY 133 PVGHKLAVWNNDIYVKIEPNLPSYRIWTGKEDIYNGITDWWYEEVFSAYSALWSP 192  
DB 159 PVGHKLAVWNNDIYVKIEPNLPSYRIWTGKEDIYNGITDWWYEEVFSAYSALWSP 218  
QY 193 NGTFLAYAFNDTEVPLIEYSFSDESIYQPKTVRVPYKAGAVNPTKFFVNTDLSLS 252  
DB 219 NGTFLAYAFNDTEVPLIEYSFSDESIYQPKTVRVPYKAGAVNPTKFFVNTDLSLS 278  
QY 253 VTNATSIQTAPASMLIGDHYLCVDTWATQBRISLQWLRRIQNTSVMDICDYDSSGRWN 312  
DB 279 VTNATSIQTAPASMLIGDHYLCVDTWATQBRISLQWLRRIQNTSVMDICDYDSSGRWN 338  
QY 313 CLVARQHLEMTTGVGRFSEPHFTLDGNSFYKIIISNREGYRHCYFQIDKDCCTFIT 372  
DB 339 CLVARQHLEMTTGVGRFSEPHFTLDGNSFYKIIISNREGYRHCYFQIDKDCCTFIT 398  
QY 373 KGTWEVIGIEALTSDYLYISNEYKMGPGGNLYKIQLSDYTKVTCLSCNLPBCOYYS 432  
DB 399 KGTWEVIGIEALTSDYLYISNEYKMGPGGNLYKIQLSDYTKVTCLSCNLPBCOYYS 458  
QY 433 VFSKBAKTYQLRCSGPGCLPLYTLHSSVNDKGLRVLEDNSALDQQLQNVQMPKCLDFII 492  
DB 459 VFSKBAKTYQLRCSGPGCLPLYTLHSSVNDKGLRVLEDNSALDQQLQNVQMPKCLDFII 518

QY 493 LNETKFWYQMLPPHFDKSKYPLLLDVYAGPCSKADTVPRLNWATVLASTENIIVASP 552  
DB 519 LNETKFWYQMLPPHFDKSKYPLLLDVYAGPCSKADTVPRLNWATVLASTENIIVASP 578  
QY 553 DGRSGYQDQKIMHAINRRLGTFFVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVVTSM 612  
DB 579 DGRSGYQDQKIMHAINRRLGTFFVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVVTSM 638  
QY 613 VLGSYGVPKCGIAVAPVSRWEYDVSVTERYMGLTPTEDNLDHYRNSTVMSRAENFKOV 672  
DB 639 VLGSYGVPKCGIAVAPVSRWEYDVSVTERYMGLTPTEDNLDHYRNSTVMSRAENFKOV 698  
QY 673 EYLLHGTADDNVHFQQAQISKALVDGVDFQAMWYTDDEHGIASSTAHQHIYTHMSHF 732  
DB 699 EYLLHGTADDNVHFQQAQISKALVDGVDFQAMWYTDDEHGIASSTAHQHIYTHMSHF 758  
QY 733 IKQCFSLP 740  
DB 759 IKQCFSLP 766

## RESULT 7

US-09-265-606-3  
; Sequence 3, Application US/09265606  
; Patent No. 6846910  
; GENERAL INFORMATION:  
; APPLICANT: Zimmermann, Rainer; Park, John E.;  
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.  
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN  
; TITLE OF INVENTION: ALPHA, AND USES THEREOF  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felfe & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/265,606  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/619,280  
; FILING DATE: 18-MARCH-1996  
; APPLICATION NUMBER: 08/230,491  
; FILING DATE: 20-APRIL-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hanson, No. 6846910man D.  
; REGISTRATION NUMBER: 30,946  
; REFERENCE/DOCKET NUMBER: LUD 5330.1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 688-9200  
; TELEFAX: (212) 838-3884  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 766 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-09-265-606-3

Query Match 98.0%; Score 3939; DB 2; Length 766;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 13 SRKTYTLDYLNKTYRLKLSLRWISDHEYLKQENNILVFNAYGNSVFLNSTDEF 72

Db 39 SRKTYTLTDYLNKNTYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVPLENSTFDEF 98  
QY 73 GHSINDYSISPDGQFILLEYNVVKWRHSYTYASYDIYDLNKRQLITEERIPNNTQVWTWS 132  
Db 99 GHSINDYSISPDGQFILLEYNVVKWRHSYTYASYDIYDLNKRQLITEERIPNNTQVWTWS 158  
QY 133 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFSAYSALWSP 192  
Db 159 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFSAYSALWSP 218  
QY 193 NGTFLAYAQFNDTEVPLLEYSFYSDLSQYPKTVRPVPKAGAVNPTVKFFVNTDLS 252  
Db 219 NGTFLAYAQFNDTEVPLLEYSFYSDLSQYPKTVRPVPKAGAVNPTVKFFVNTDLS 278  
QY 253 VTNATSIQITAPASMLIGDHYLCVDTWATQERISLQWLRRIONYSVMDICDYDESSGRWN 312  
Db 279 VTNATSIQITAPASMLIGDHYLCVDTWATQERISLQWLRRIONYSVMDICDYDESSGRWN 338  
QY 313 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFIT 372  
Db 339 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFIT 398  
QY 373 KGTWEVIGIEALTSDYLYIISNEYKMPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS 432  
Db 399 KGTWEVIGIEALTSDYLYIISNEYKMPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS 458  
QY 433 VSFSKEAKYQYLRCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFII 492  
Db 459 VSFSKEAKYQYLRCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFII 518  
QY 493 LNETKFWYQMLPPHFDKSKYPLLLDYYVAGPCSKADTVFRLNWTYLASTENIIVASF 552  
Db 519 LNETKFWYQMLPPHFDKSKYPLLLDYYVAGPCSKADTVFRLNWTYLASTENIIVASF 578  
QY 553 DGRSGYGQDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGYVTSM 612  
Db 579 DGRSGYGQDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGYVTSM 638  
QY 613 VLGSGGVFKCGIAPVSRWEYDSYTYTERYMGLTPTEDNLDHYRNSTVMSRAENFKQV 672  
Db 639 VLGSGGVFKCGIAPVSRWEYDSYTYTERYMGLTPTEDNLDHYRNSTVMSRAENFKQV 698  
QY 673 EYLLIHGTADDNVHFQOQSAQISKALVDGVDFQAMWYTDDEHGIASSTAHOIYTHMSHF 732  
Db 699 EYLLIHGTADDNVHFQOQSAQISKALVDGVDFQAMWYTDDEHGIASSTAHOIYTHMSHF 758  
QY 733 IKQCFSLP 740  
Db 759 IKQCFSLP 766  
RESULT 8  
US-09-016-10450  
; Sequence 10450, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10450  
; LENGTH: 775  
; TYPE: PR1

; ORGANISM: Human  
US-09-949-016-10450  
Query Match 98.0%; Score 3939; DB 2; Length 775;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 728; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 13 SRKTYTLTDYLNKNTYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVPLENSTFDEF 72  
Db 48 SRKTYTLTDYLNKNTYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVPLENSTFDEF 107  
QY 73 GHSINDYSISPDGQFILLEYNVVKWRHSYTYASYDIYDLNKRQLITEERIPNNTQVWTWS 132  
Db 108 GHSINDYSISPDGQFILLEYNVVKWRHSYTYASYDIYDLNKRQLITEERIPNNTQVWTWS 167  
QY 133 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFSAYSALWSP 192  
Db 168 PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFSAYSALWSP 227  
QY 193 NGTFLAYAQFNDTEVPLLEYSFYSDLSQYPKTVRPVPKAGAVNPTVKFFVNTDLS 252  
Db 228 NGTFLAYAQFNDTEVPLLEYSFYSDLSQYPKTVRPVPKAGAVNPTVKFFVNTDLS 287  
QY 253 VTNATSIQITAPASMLIGDHYLCVDTWATQERISLQWLRRIONYSVMDICDYDESSGRWN 312  
Db 288 VTNATSIQITAPASMLIGDHYLCVDTWATQERISLQWLRRIONYSVMDICDYDESSGRWN 347  
QY 313 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFIT 372  
Db 348 CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFIT 407  
QY 373 KGTWEVIGIEALTSDYLYIISNEYKMPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS 432  
Db 408 KGTWEVIGIEALTSDYLYIISNEYKMPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS 467  
QY 433 VSFSKEAKYQYLRCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFII 492  
Db 468 VSFSKEAKYQYLRCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFII 527  
QY 493 LNETKFWYQMLPPHFDKSKYPLLLDYYVAGPCSKADTVFRLNWTYLASTENIIVASF 552  
Db 528 LNETKFWYQMLPPHFDKSKYPLLLDYYVAGPCSKADTVFRLNWTYLASTENIIVASF 587  
QY 553 DGRSGYGQDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGYVTSM 612  
Db 588 DGRSGYGQDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGYVTSM 647  
QY 613 VLGSGGVFKCGIAPVSRWEYDSYTYTERYMGLTPTEDNLDHYRNSTVMSRAENFKQV 672  
Db 648 VLGSGGVFKCGIAPVSRWEYDSYTYTERYMGLTPTEDNLDHYRNSTVMSRAENFKQV 707  
QY 673 EYLLIHGTADDNVHFQOQSAQISKALVDGVDFQAMWYTDDEHGIASSTAHOIYTHMSHF 732  
Db 708 EYLLIHGTADDNVHFQOQSAQISKALVDGVDFQAMWYTDDEHGIASSTAHOIYTHMSHF 767  
QY 733 IKQCFSLP 740  
Db 768 IKQCFSLP 775  
RESULT 9  
US-10-002-593-6  
; Sequence 6, Application US/10002593  
; Patent No. 6586198  
; GENERAL INFORMATION:  
; APPLICANT: Vanderbilt University  
; APPLICANT: Brown, Nancy J.  
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOGENESIS CONVERTING  
; INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA  
; FILE REFERENCE: US/10/002,593  
; CURRENT APPLICATION NUMBER: US/10-002,593  
; CURRENT FILING DATE: 2001-10-31  
; PRIOR APPLICATION NUMBER: 60/244,524

; PRIOR FILING DATE: 2000-10-31  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 6  
 ; LENGTH: 766  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-002-593-6

Query Match		97.8%	Score 3933;	DB 2;	Length 766;
Best Local Similarity		99.9%	Pred. No. 0;		
Matches	727;	Conservative	0;	Mismatches	1; Indels 0; Gaps 0;

  

QY	13	SRKTYTLTDYLNKNTYRLKLSLRWISDHEYLKQENNILVFNABYGNSSVFLNSTPDEF	72
Db	39	SRKTYTLTDYLNKNTYRLKLSLRWISDHEYLKQENNILVFNABYGNSSVFLNSTPDEF	98
QY	73	GHSINDYSISPDGQFILLLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVWS	132
Db	99	GHSINDYSISPDGQFILLLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVWS	158
QY	133	PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWNSP	192
Db	159	PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWNSP	218
QY	193	NGTFLAYAQFNDTEVPLIEYSFYSDLSQYPKTVRPYPKAGAVNPTVKFFVNTDSLSS	252
Db	219	NGTFLAYAQFNDTEVPLIEYSFYSDLSQYPKTVRPYPKAGAVNPTVKFFVNTDSLSS	278
QY	253	VTNATSIQITAPASMLIGDHYLCVDTWATQERISLOWLRRIQNYSVMDICDYDESSGRWN	312
Db	279	VTNATSIQITAPASMLIGDHYLCVDTWATQERISLOWLRRIQNYSVMDICDYDESSGRWN	338
QY	313	CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEGYRHI CYFQIDKXDCETIT	372
Db	339	CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEGYRHI CYFQIDKXDCETIT	398
QY	373	KGTWEVIGIEALTSDLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS	432
Db	399	KGTWEVIGIEALTSDLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS	458
QY	433	VFSKEAKYQLRCSGPGPLPYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKCLDPII	492
Db	459	VFSKEAKYQLRCSGPGPLPYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKCLDPII	518
QY	493	LNETHFWQMLPDPHFKSKYKYPILLDVYAGPCQKADTVPRLNWATYLASTENIIIVASF	552
Db	519	LNETHFWQMLPDPHFKSKYKYPILLDVYAGPCQKADTVPRLNWATYLASTENIIIVASF	578
QY	553	DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGYVTSM	612
Db	579	DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGYVTSM	638
QY	613	VLGSGGVFKGIAVAPVSRWEYDYSVYTERYMGLEPTPEDNLDHYRNSVMSRAENFKQV	672
Db	639	VLGSGGVFKGIAVAPVSRWEYDYSVYTERYMGLEPTPEDNLDHYRNSVMSRAENFKQV	698
QY	673	EYLLIHGTADDNVHFQOQSAQISKALVDVGVDFQAMWYTDDEHGIIASSTAHQHIYTHMSHF	732
Db	699	EYLLIHGTADDNVHFQOQSAQISKALVDVGVDFQAMWYTDDEHGIIASSTAHQHIYTHMSHF	758
QY	733	IKQCFSLP 740	
Db	759	IKQCFSLP 766	

RESULT 10  
 US-09-949-016-6146  
 ; Sequence 6146. Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 6146  
 ; LENGTH: 766  
 ; TYPE: PRT  
 ; ORGANISM: Human  
 US-09-949-016-6146

Query Match		97.8%	Score 3933;	DB 2;	Length 766;
Best Local Similarity		99.9%	Pred. No. 0;		
Matches	727;	Conservative	0;	Mismatches	1; Indels 0; Gaps 0;

  

QY	13	SRKTYTLTDYLNKNTYRLKLSLRWISDHEYLKQENNILVFNABYGNSSVFLNSTPDEF	72
Db	39	SRKTYTLTDYLNKNTYRLKLSLRWISDHEYLKQENNILVFNABYGNSSVFLNSTPDEF	98
QY	73	GHSINDYSISPDGQFILLLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVWS	132
Db	99	GHSINDYSISPDGQFILLLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVWS	158
QY	133	PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWNSP	192
Db	159	PVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWNSP	218
QY	193	NGTFLAYAQFNDTEVPLIEYSFYSDLSQYPKTVRPYPKAGAVNPTVKFFVNTDSLSS	252
Db	219	NGTFLAYAQFNDTEVPLIEYSFYSDLSQYPKTVRPYPKAGAVNPTVKFFVNTDSLSS	278
QY	253	VTNATSIQITAPASMLIGDHYLCVDTWATQERISLOWLRRIQNYSVMDICDYDESSGRWN	312
Db	279	VTNATSIQITAPASMLIGDHYLCVDTWATQERISLOWLRRIQNYSVMDICDYDESSGRWN	338
QY	313	CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEGYRHI CYFQIDKXDCETIT	372
Db	339	CLVARQHIEMSTTGWGRFRPSEPHFTLDGNSFYKIIISNEGYRHI CYFQIDKXDCETIT	398
QY	373	KGTWEVIGIEALTSDLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS	432
Db	399	KGTWEVIGIEALTSDLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYS	458
QY	433	VFSKEAKYQLRCSGPGPLPYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKCLDPII	492
Db	459	VFSKEAKYQLRCSGPGPLPYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKCLDPII	518
QY	493	LNETHFWQMLPDPHFKSKYKYPILLDVYAGPCQKADTVPRLNWATYLASTENIIIVASF	552
Db	519	LNETHFWQMLPDPHFKSKYKYPILLDVYAGPCQKADTVPRLNWATYLASTENIIIVASF	578
QY	553	DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGYVTSM	612
Db	579	DGRSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGYVTSM	638
QY	613	VLGSGGVFKGIAVAPVSRWEYDYSVYTERYMGLEPTPEDNLDHYRNSVMSRAENFKQV	672
Db	639	VLGSGGVFKGIAVAPVSRWEYDYSVYTERYMGLEPTPEDNLDHYRNSVMSRAENFKQV	698
QY	673	EYLLIHGTADDNVHFQOQSAQISKALVDVGVDFQAMWYTDDEHGIIASSTAHQHIYTHMSHF	732
Db	699	EYLLIHGTADDNVHFQOQSAQISKALVDVGVDFQAMWYTDDEHGIIASSTAHQHIYTHMSHF	758
QY	733	IKQCFSLP 740	
Db	759	IKQCFSLP 766	

RESULT 11  
US-10-423-714-6  
; Sequence 6, Application US/10423714  
; Patent No. 6887679  
; GENERAL INFORMATION:  
; APPLICANT: Vanderbilt University  
; APPLICANT: Brown, Nancy J.  
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTIN  
; TITLE OF INVENTION: INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIODEMA  
; FILE REFERENCE: Acty Docket No. 6887679 1242/48/2  
; CURRENT APPLICATION NUMBER: US/10/423,714  
; CURRENT FILING DATE: 2003-04-25  
; PRIOR APPLICATION NUMBER: 60/244,524  
; PRIOR FILING DATE: 2000-10-31  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 766  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-423-714-6

Query Match 97.8%; Score 3933; DB 2; Length 766;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 727; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 SRKTYTLDTYLNKNTYRLKLYSLRWISDHEYLKQENNLVFNAYGNSSVFLNSTFDEF 72  
DB 39 SRKTYTLDTYLNKNTYRLKLYSLRWISDHEYLKQENNLVFNAYGNSSVFLNSTFDEF 98

QY 73 GHSINDYSISPDQGFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQVTTWS 132  
DB 99 GHSINDYSISPDQGFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQVTTWS 158

QY 133 PVGHKLAYVNNNDIYVKIEPNLPSYRIITWTGKEDIYNGITDWTVEEVEFSAYSAWWS 192  
DB 159 PVGHKLAYVNNNDIYVKIEPNLPSYRIITWTGKEDIYNGITDWTVEEVEFSAYSAWWS 218

QY 193 NGTFLAYAQFNDTEVPLIEYSFYSDESLOYPKTVRPVYPKAGAVNPTVKFFVNTDLS 252  
DB 219 NGTFLAYAQFNDTEVPLIEYSFYSDESLOYPKTVRPVYPKAGAVNPTVKFFVNTDLS 278

QY 253 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLOWLRRIQNSYVMDICDYDESSGRWN 312  
DB 279 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLOWLRRIQNSYVMDICDYDESSGRWN 338

QY 313 CLVAROHIE MSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCCTFIT 372  
DB 339 CLVAROHIE MSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCCTFIT 398

QY 373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELAPERCOYYS 432  
DB 399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELAPERCOYYS 458

QY 433 VSFSKEAKYQYLRCSGPGPLPYTLHSSVNDKGLRVLEEDNSALDKMLQNVQMPSSKCLDFII 492  
DB 459 VSFSKEAKYQYLRCSGPGPLPYTLHSSVNDKGLRVLEEDNSALDKMLQNVQMPSSKCLDFII 518

QY 493 LNETKFWYQMLPPHFDKSKYPLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASF 552  
DB 519 LNETKFWYQMLPPHFDKSKYPLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASF 578

QY 553 DGRGSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGYVTSM 612  
DB 579 DGRGSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGYVTSM 638

QY 613 VLGSQGVKFCGKIAPVSRWEYDYSYTERYNGLPTPBEDNLDPYRNSVTMGRAENFKOV 672  
DB 639 VLGSQGVKFCGKIAPVSRWEYDYSYTERYNGLPTPBEDNLDPYRNSVTMGRAENFKOV 698

QY 673 EYLLIHGTADDNVHFQSSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHF 732

Db 699 EYLLIHGTADDNVHFQSSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHF 758

QY 733 IKQCFSLP 740  
Db 759 IKQCFSLP 766

RESULT 12  
US-09-518-550-27  
; Sequence 27, Application US/09518550  
; Patent No. 6875851  
; GENERAL INFORMATION:  
; APPLICANT: POTEMPA, James  
; APPLICANT: BANBULA, Agnieszka  
; TITLE OF INVENTION: PROLYL PEPTIDASES AND METHODS OF USE  
; FILE REFERENCE: 235,00190101  
; CURRENT APPLICATION NUMBER: US/09/518,550  
; CURRENT FILING DATE: 2000-03-03  
; PRIOR APPLICATION NUMBER: 60/123,148  
; PRIOR FILING DATE: 1999-03-05  
; PRIOR APPLICATION NUMBER: PCT/US00/05551  
; PRIOR FILING DATE: 2000-03-03  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 27  
; LENGTH: 766  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-518-550-27

Query Match 97.7%; Score 3929; DB 2; Length 766;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 726; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 13 SRKTYTLDTYLNKNTYRLKLYSLRWISDHEYLKQENNLVFNAYGNSSVFLNSTFDEF 72  
DB 39 SRKTYTLDTYLNKNTYRLKLYSLRWISDHEYLKQENNLVFNAYGNSSVFLNSTFDEF 98

QY 73 GHSINDYSISPDQGFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQVTTWS 132  
DB 99 GHSINDYSISPDQGFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQVTTWS 158

QY 133 PVGHKLAYVNNNDIYVKIEPNLPSYRIITWTGKEDIYNGITDWTVEEVEFSAYSAWWS 192  
DB 159 PVGHKLAYVNNNDIYVKIEPNLPSYRIITWTGKEDIYNGITDWTVEEVEFSAYSAWWS 218

QY 193 NGTFLAYAQFNDTEVPLIEYSFYSDESLOYPKTVRPVYPKAGAVNPTVKFFVNTDLS 252  
DB 219 NGTFLAYAQFNDTEVPLIEYSFYSDESLOYPKTVRPVYPKAGAVNPTVKFFVNTDLS 278

QY 253 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLOWLRRIQNSYVMDICDYDESSGRWN 312  
DB 279 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLOWLRRIQNSYVMDICDYDESSGRWN 338

QY 313 CLVAROHIE MSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCCTFIT 372  
DB 339 CLVAROHIE MSTTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDCCTFIT 398

QY 373 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELAPERCOYYS 432  
DB 399 KGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELAPERCOYYS 458

QY 433 VSFSKEAKYQYLRCSGPGPLPYTLHSSVNDKGLRVLEEDNSALDKMLQNVQMPSSKCLDFII 492  
DB 459 VSFSKEAKYQYLRCSGPGPLPYTLHSSVNDKGLRVLEEDNSALDKMLQNVQMPSSKCLDFII 518

QY 493 LNETKFWYQMLPPHFDKSKYPLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASF 552  
DB 519 LNETKFWYQMLPPHFDKSKYPLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASF 578

QY 553 DGRGSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGYVTSM 612



Db 579 DGRSGYQGDIMHAINRLTFFVEDQIEAARQFSKMGFVDNKRRIAIWGSYGYVTSM 638  
Qy 613 VLGSAGVFKGIAVAPVSRWEYDVSVTERYMGLPPTPEDNLDHYRSTVMSRAENFAQV 672  
Db 639 VLGSAGVFKGIAVAPVSRWEYDVSVTERYMGLPPTPEDNLDHYRSTVMSRAENFAQV 698  
Qy 673 EYLLIHGTADNVHFQSSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHQHITHMSHF 732  
Db 699 EYLLIHGTADNVHFQSSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHQHITHMSHF 758  
Qy 733 IKQCFSLP 740  
Db 759 IKQCFSLP 766

RESULT 13

PCT-US93-07923-11  
; Sequence 11, Application PC/TUS9307923  
; GENERAL INFORMATION:  
; APPLICANT: Morimoto, Chikao  
; APPLICANT: Schlossman, Stuart F.  
; APPLICANT: Tanaka, Toshiaki  
; TITLE OF INVENTION: HUMAN CD26 AND METHODS FOR USE  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 50Z or 55SX  
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
; SOFTWARE: WordPerfect (Version 5.0)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/07923  
; FILING DATE: 19930819  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/934,162  
; FILING DATE: 21-AUG-1992  
; APPLICATION NUMBER: 07/832,211  
; FILING DATE: 06-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frazer, Janis K.  
; REGISTRATION NUMBER: 34,819  
; REFERENCE/DOCKET NUMBER: 00530/055002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 593  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
PCT-US93-07923-11

Query Match 74.9%; Score 3010; DB 4; Length 593;  
Best Local Similarity 100.0%; Pred. No. 4.9e-256;  
Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 13 SRKTYTLTDYLNKTYRLKLSLRWISDHEYLKQENNILVFNASVGNSSVFLNSTDEF 72  
Db 39 SRKTYTLTDYLNKTYRLKLSLRWISDHEYLKQENNILVFNASVGNSSVFLNSTDEF 98  
Qy 73 GHSINDYSISPDGPFILLEYNYVQWRHSYTSYDIDLNKRQLITERIPNNTQWTVWS 132  
Db 99 GHSINDYSISPDGPFILLEYNYVQWRHSYTSYDIDLNKRQLITERIPNNTQWTVWS 158

Qy 133 PVGHKLAYWNNDIYVKIEPNLPSYRITWTGKBDIIYNGITDWWYEEVFSAYSALWMSF 192  
Db 159 PVGHKLAYWNNDIYVKIEPNLPSYRITWTGKBDIIYNGITDWWYEEVFSAYSALWMSF 218  
Qy 193 NGTFPLAYAQNDTEVPLIEYSFYSDESLOYPKTVRVPYPKAGAVNPTVKPVVNTDLSL 252  
Db 219 NGTFPLAYAQNDTEVPLIEYSFYSDESLOYPKTVRVPYPKAGAVNPTVKPVVNTDLSL 278  
Qy 253 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLOWLRRIQNSVMDICDYDESSGRWN 312  
Db 279 VTNATSIQITAPASMLIGDHYLCDVTWATQERISLOWLRRIQNSVMDICDYDESSGRWN 338  
Qy 313 CLVARQHIEMSTTGWGRPRPSPHFTLDGNSFYKIIISNEEGYRHICYFQIDKDKCFTIT 372  
Db 339 CLVARQHIEMSTTGWGRPRPSPHFTLDGNSFYKIIISNEEGYRHICYFQIDKDKCFTIT 398  
Qy 373 KGTWEVIGIEALTSDYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYS 432  
Db 399 KGTWEVIGIEALTSDYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYS 458  
Qy 433 VSPSKBAKYQLRCGPGPLPLYTLHSSVNDKGLRVLEDNSALDQMLQNVOMPSKCLDPII 492  
Db 459 VSPSKBAKYQLRCGPGPLPLYTLHSSVNDKGLRVLEDNSALDQMLQNVOMPSKCLDPII 518  
Qy 493 LNETKFWYQMLPPHFDKSKYPLLLDVYAGPCSKADTVFRLNWTYLASTENIIIVASF 552  
Db 519 LNETKFWYQMLPPHFDKSKYPLLLDVYAGPCSKADTVFRLNWTYLASTENIIIVASF 578  
Qy 553 DGRSGYQGDIMHIA 567  
Db 579 DGRSGYQGDIMHIA 593  
RESULT 14  
US-08-230-491A-2  
; Sequence 2, Application US/08230491A  
; Patent No. 5587299  
; GENERAL INFORMATION:  
; APPLICANT: Rettig, Wolfgang J.; Scanlan, Matthew J.;  
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING FOR  
; TITLE OF INVENTION: FIBROBLAST ACTIVATION PROTEIN ' AND USES  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FELPE & LYNCH  
; STREET: 805 THIRD AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: DISKETTE 3.5 inch 1.2 MB STORAGE  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: WORDPERFECT - ASC II  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/230,491A  
; FILING DATE: 20-APRIL-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hanson, No. 5587299man D.  
; REGISTRATION NUMBER: 30,946  
; REFERENCE/DOCKET NUMBER: LUD 330  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 688-9200  
; TELEFAX: (212) 838-3884  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 760 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

US-08-230-491A-2

Query Match 53.9%; Score 2168; DB 1; Length 760;  
Best Local Similarity 52.5%; Pred. No. 9.9e-182;  
Matches 389; Conservative 138; Mismatches 200; Indels 14; Gaps 7;  
QY 3 PGGSHHHHSHKRTVLTDLKNTYRLKLYSLRWISDHELYKO-ENNILVFNAEYGNSS 61  
DB 28 PSRVHSEBNTWRLTLKDLINGTFSYKTFPFWNLSGQYLFHQSDANNVILNIETGQSY 87  
QY 62 VFLENSTDFEGHSIN--DYSISPDGQFILLBYNYVKKWRHSYTSYDIYDLNKRQLITE 119  
DB 88 TILSNRTM----KSNVSNYGLSPDRQFVYLESYKLRYSYATYIYIDLNGEFVRG 143  
QY 120 ERIPNNTQWVTSVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEE 179  
DB 144 NELPRPIQYLCWSPGVSKLAYVYQNNIYKQRPDPPQIIFNGRENKIFNGIPDWVYEE 203  
QY 180 EVFSAYSALWSPNGTFLAYAOFNDETVPLIEYFSYSDSLOYPKTVRPVYPKAGAVNPT 239  
DB 204 EMLPTKYALWSPNGKFLAYAEFNDKDIPIVAYSYGDE--QYPTINIPYPKAGAKNPV 261  
QY 240 VKFVVNTDSLSSVNTNATSIQITAPASMLIGHYLCVDTWATQERISLQWLRRIQNSVM 299  
DB 262 VRIFIIDTTYPAVGPQ---EVPVPAMIASDYFSPSWLTWVTDERVKLQWLRVQNVSVL 318  
QY 300 DICVDDESSGRNCLVARQHIEMSTTGWVGRFRPSEPHEFTLDGNSFYKIIISNEEGYRHC 359  
DB 319 SICDFREDQWDCPKTQHEIESRTGWAGFFVSRPVFSYDAISYKIFSDKGYGKIH 378  
QY 360 YFQIDKDCOTFTTKGTWVIGIEALTSYLYISNEYKMGPGRNLYKIQLSDY-TKVTVC 418  
DB 379 YIKDTVENAIQITSGKWEAINFRVTQDSLFSYSSNEFEYPPGRNIYRISIGSYPPSKKC 438  
QY 419 LSCELNPERCOYYSVSFSKEAKYQLRCSGQGLPLYTLHSSVNDKGLRVLEDNSALDKML 478  
DB 439 VTCHLRKERCQYYTASFSDYAKYALVCYGPPISTLHDGRTDQEIKEILEENKLENAL 498  
QY 479 QNVQPSKKLDPIILNETKFWQMILPHFDKSKYKYPILLDYGAGCSQKADTVFRLNWA 538  
DB 499 KNIQLPKEIEKKLEVDLTLAYKMLLPQFDRSKYKYPILLIYVGGPSCQSRSVFAVNW 558  
QY 539 TYLASTENIIVASFDRGSGYGQDKIMHAINRRLGTFTFEVDQIEAARQFSKMGFVDNKR 598  
DB 559 SYLASKEGMVIALVDGRGTAFQDGKLLYAVYRKLGVVEVEDQITAVRKFIENGFIDEKRI 618  
QY 599 AIWGSYGGYVTSMLVSGSGVFKGCIAPVAPSRWEYDSVYTERYMGLPPTEDNLDHYR 658  
DB 619 AIWGSYGGYVSSLASGTGLFKGCIAPVAPSSWEYASVYTERFMGLPTKDDNLEHYK 678  
QY 659 NSTVMSRAENPKQVEYLLHGTADNNVHPQQAQISKALVDVGVDFQAMWYTDDEHGIAS 718  
DB 679 NSTVMARAEYFRNVYLLHGTADNNVHPQNSAQIAKALVNAQVDFQAMWYSDQNHGL-S 737  
QY 719 STAHOHYTHMSHFITKQCFSL 739  
DB 738 GUSTNRLYTHMTHFLKQCFSL 758

RESULT 15  
US-08-619-280A-2  
Sequence 2, Application US/08619280A  
Patent No. 5767242  
GENERAL INFORMATION:  
APPLICANT: Zimmermann, Rainer; Park, John E.;  
APPLICANT: Rettig, Wolfgang; Old, Lloyd J.  
TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN  
TITLE OF INVENTION: ALPHA, AND USES THEREOF  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESS: Felife & Lynch  
STREET: 805 Third Avenue  
CITY: New York City

STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/619,280A  
FILING DATE: 18-MARCH-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/230,491  
FILING DATE: 20-APRIL-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 5767242man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5330.1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 760 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-619-280A-2

Query Match 53.9%; Score 2168; DB 1; Length 760;  
Best Local Similarity 52.5%; Pred. No. 9.9e-182;  
Matches 389; Conservative 138; Mismatches 200; Indels 14; Gaps 7;  
QY 3 PGGSHHHHSHKRTVLTDLKNTYRLKLYSLRWISDHELYKO-ENNILVFNAEYGNSS 61  
DB 28 PSRVHSEBNTWRLTLKDLINGTFSYKTFPFWNLSGQYLFHQSDANNVILNIETGQSY 87  
QY 62 VFLENSTDFEGHSIN--DYSISPDGQFILLBYNYVKKWRHSYTSYDIYDLNKRQLITE 119  
DB 88 TILSNRTM----KSNVSNYGLSPDRQFVYLESYKLRYSYATYIYIDLNGEFVRG 143  
QY 120 ERIPNNTQWVTSVGHKLAYVWNNDIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEE 179  
DB 144 NELPRPIQYLCWSPGVSKLAYVYQNNIYKQRPDPPQIIFNGRENKIFNGIPDWVYEE 203  
QY 180 EVFSAYSALWSPNGTFLAYAOFNDETVPLIEYFSYSDSLOYPKTVRPVYPKAGAVNPT 239  
DB 204 EMLPTKYALWSPNGKFLAYAEFNDKDIPIVAYSYGDE--QYPTINIPYPKAGAKNPV 261  
QY 240 VKFVVNTDSLSSVNTNATSIQITAPASMLIGHYLCVDTWATQERISLQWLRRIQNSVM 299  
DB 262 VRIFIIDTTYPAVGPQ---EVPVPAMIASDYFSPSWLTWVTDERVKLQWLRVQNVSVL 318  
QY 300 DICVDDESSGRNCLVARQHIEMSTTGWVGRFRPSEPHEFTLDGNSFYKIIISNEEGYRHC 359  
DB 319 SICDFREDQWDCPKTQHEIESRTGWAGFFVSRPVFSYDAISYKIFSDKGYGKIH 378  
QY 360 YFQIDKDCOTFTTKGTWVIGIEALTSYLYISNEYKMGPGRNLYKIQLSDY-TKVTVC 418  
DB 379 YIKDTVENAIQITSGKWEAINFRVTQDSLFSYSSNEFEYPPGRNIYRISIGSYPPSKKC 438  
QY 419 LSCELNPERCOYYSVSFSKEAKYQLRCSGQGLPLYTLHSSVNDKGLRVLEDNSALDKML 478  
DB 439 VTCHLRKERCQYYTASFSDYAKYALVCYGPPISTLHDGRTDQEIKEILEENKLENAL 498  
QY 479 QNVQPSKKLDPIILNETKFWQMILPHFDKSKYKYPILLDYGAGCSQKADTVFRLNWA 538  
DB 499 KNIQLPKEIEKKLEVDLTLAYKMLLPQFDRSKYKYPILLIYVGGPSCQSRSVFAVNW 558  
QY 539 TYLASTENIIVASFDRGSGYGQDKIMHAINRRLGTFTFEVDQIEAARQFSKMGFVDNKR 598  
DB 559 SYLASKEGMVIALVDGRGTAFQDGKLLYAVYRKLGVVEVEDQITAVRKFIENGFIDEKRI 618

Qy	599	AIWGSYGGYVTSMLGSGGVFKGCIAPVSRWEYDVSVYTERYMGLETPEDNLDHYR	658
Db	619	AIWGSYGGYVSSLALASGTGLFKCIIAPVSSWEYASVYTERFMGLPTKDDNLEHYK	678
Qy	659	NSTVMSRAENFKQVEYLLIHGTADDNVHFPQSAQISKALVDVGVDFOAMWYTDDEHGIAS	718
Db	679	NSTVWARAEYFRVDYLLIHGTADDNVHFPQSAQIAKALVNAQVDFQAMWYSDQNHGL-S	737
Qy	719	STAHOIYTHMSHFQKCFSL	739
Db	738	GLSTNHLTYTHMTHFLKQCFSL	758

Search completed: February 15, 2006, 13:16:21  
Job time : 51 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2006, 04:32:36 ; Search time 8805 Seconds  
(without alignments)  
11605.109 Million cell updates/sec

Title: US-10-659-055-2  
Perfect score: 2184  
Sequence: 1 agtcgcaaaactacactct.....aacaaatgttcttcttacct 2184

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:

- 1: gb\_est1.\*
- 2: gb\_est2.\*
- 3: gb\_est3.\*
- 4: gb\_hc.\*
- 5: gb\_est4.\*
- 6: gb\_est5.\*
- 7: gb\_est6.\*
- 8: gb\_est7.\*
- 9: gb\_ges1.\*
- 10: gb\_ges2.\*
- 11: gb\_ges3.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2153.6	98.6	3411	4	CR860150 Pongo pyg
2	1677.6	76.8	3289	4	AK085370 Mus muscu
3	1549.4	70.9	2193	10	AY418368 Homo sapi
4	1536.8	70.4	2193	10	AY418369 Pan trogl
5	1237	56.6	2191	10	AY418370 Mus muscu
6	1041.2	47.7	1081	1	AL554569 AL554569
7	865	39.6	987	5	BX399854 BX399854
8	829.2	38.0	886	5	BX348426 BX348426
9	818	37.5	1014	1	AL548644 AL548644
10	803.2	36.8	1022	1	AL541949 AL541949
11	795.8	36.4	1007	1	AL544456 AL544456
12	768.4	35.2	846	1	AU122649 AU122649
13	722	33.1	900	5	BQ937942 AGENCOURT
14	721	33.0	897	1	AL553051 AL553051
15	712.6	32.6	791	1	AU139683 AU139683
16	704	32.2	731	6	CB957930 AGENCOURT
17	701.4	32.1	741	5	BX496051 DKFZp779P
18	695.8	31.9	904	5	BX397280 BX397280
19	683	31.3	773	1	AU139062 AU139062
20	682.6	31.3	2277	10	AY402291 Homo sapi
21	678.4	31.1	2277	10	AY402292 Pan trogl
22	675.6	30.9	887	1	AL550603 AL550603

23	671.4	30.7	1087	6	CF109988	CF109988 Shultzomi
24	664.8	30.4	876	6	CD516785	CD516785 AGENCOURT
25	658.4	30.1	2280	10	AY402293	AY402293 Mus muscu
26	646.2	29.6	742	7	CK473401	CK473401 AGENCOURT
27	642.6	29.4	692	6	CD613465	CD613465 551138877
28	636.8	29.2	640	6	CB211864	CB211864 NISC_nq09
29	628.6	28.8	717	7	CN786863	CN786863 4120921 B
30	627.4	28.7	733	3	BP448857	BP448857 BP448857
31	620.2	28.4	772	3	BI560727	BI560727 603254767
32	619.4	28.4	642	6	CB265788	CB265788 1004693 H
33	618.6	28.3	915	5	BQ891954	BQ891954 AGENCOURT
34	600.4	27.5	930	2	BF982962	BF982962 602306363
35	591	27.1	826	7	CR769129	CR769129 DKFZp4696
36	582	26.6	583	3	BP334755	BP334755 BP334755
37	581	26.6	581	3	BP276380	BP276380 BP276380
38	575.2	26.3	913	3	BP442613	BP442613 BP442613
39	572	26.2	697	2	EG401908	EG401908 602466614
40	569.4	26.1	571	3	BF292813	BF292813 BP292813
41	569.4	26.1	609	1	AV693673	AV693673 AV693673
42	567	26.0	579	6	CD613469	CD613469 56059761J
43	566.8	26.0	583	3	BP276318	BP276318 BP276318
44	563.2	25.8	582	3	BP327817	BP327817 BP327817
45	560.2	25.7	771	5	EX392649	EX392649 EX392649

## ALIGNMENTS

RESULT 1	CR860150	Pongo pygmaeus mRNA	3411 bp	mRNA	linear	HTC 12-NOV-2004
LOCUS	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
DEFINITION	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
ACCESSION	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
VERSION	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
KEYWORDS	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
SOURCE	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
ORGANISM	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
REFERENCE	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
AUTHORS	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
CONSRMT	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
TITLE	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
JOURNAL	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
COMMENT	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
FEATURES	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
source	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
location/Qualifiers	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
1. 3411	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
/organism="Pongo pygmaeus"	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
/mol_type="mRNA"	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
/db_xref="taxon:9600"	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
/clone="DKFZp469P1419"	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
/class_type="kidney"	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
/clone_lib="469 (synonym: pkid1). Vector pSport1_Sfi; host DH10B; sites SfilA + SfilB"	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
/dev stage="adult"	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
/note="diipeptidyl peptidase IV (Homo sapiens)"	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
1. 3411	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
/gene="DKFZp469P1419"	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
63. .2360	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)
/gene="DKFZp469P1419"	CR860150	CR860150	3411 bp	CR860150	CR860150	(from clone DKFZp469P1419)

```
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH92293.1"
/db_xref="GI:55731157"
/translation="MKTPKVLGLGAAALVTIITVPVLLNKGNDATADSRKTYTL
TDLYKNTYRLKLYLSRWISDHYLYKQBNILVFNABYCNISAFLENSHPTDFEFGHSIN
DYSISPDGOFILLEYVYKWRHSVTASDVIDLNKROLITEDRIPNNTQWTVSPVG
HKLAVWNDIIVYKLEPNLPSHRIITWTKEDIIVNGITDMVYEEVPSAYSLWSPN
GTFLAYQNDTEVLLIYFSDYSDLSQPKTVRPVYPKAGAVNTPVFPVNTDLSL
SVTNATSQITAPASMLIGDHYLCDVATWATQRI SIQWLRRIQNTSVMDICDYDESSG
RWNCILVAQHIEMSTGWGRFPSEPHFTSDGNSFYKII SNEEGYRHICYPQIDKKD
CTFIKTGWEFGEIGALISDYLYISNEYKMGPNRNLKYIQLSDYTKVTCISCLRN
BRQYDYSFSEKEAYOLRCSGPLPYLYTLHSSVNDKGPRLVEDNSALDKMLQVOM
PSKLDLPILINETKFWYQMLPPHPDKSKYPLLDVAVGPCSQKADPFRVLNWTYL
ASTENIIVASDFGRSGVGQDKMHAIRRLGTFEVBDQIEAARQPSFKMGFVNDKRIA
IHWISYGGTVSNVLGSSGVFKGIAVAPSVRSWYIYDSVFTYRTMGLPTPEDNLIDHY
RNSTVMSRAENFKQVEYLLIHGTADDNVHFOQSAQISKALVDAGVDFQAMWYTDSDHG
IASSTAHOHIYTHMSHFIKQCFSLP"
```

ORIGIN

```
Query Match      98.6%; Score 2153.6; DB 4; Length 3411;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 2165; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY      1 AGTCGCAGAACTTACACTCTAACTGATTACTTAAAAAATCTATTAGACTGAAGTTATAC 60
DB      174 AGTCGCAGAACTTACACTCTAACTGATTACTTAAAAAATCTATTAGACTGAAGTTATAC 233

QY      61 TCCTTAAGAATGGAATTTAGATCATGAATATCTCTACAAACAAAGAAATATATCTTGGTA 120
DB      234 TCCTTAAGAATGGAATTTAGATCATGAATATCTCTACAAACAAAGAAATATATCTTGGTA 293

QY      121 TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 180
DB      294 TTCAATGCTGAATATGGAACAGCTCAGCTTTCTTGGAGAACAGTACATTTGATGAGTTT 353

QY      181 GGNCAATCTCATCAATGATTATCAATATCTCTGATGGGCGAGTTATCTCTTAGAATAC 240
DB      354 GGCATCTCTCATCAATGATTATCAATATCTCTGATGGGCGAGTTATCTCTTAGAATAC 413

QY      241 AACTAGCTGAAGCAATGGAGGCAATCTCTACAGCTTCTATGACATTTTATGATTTAAAT 300
DB      414 AACTATGTGAAGCAATGGAGGCAATCTCTACAGCTTCTATGACATTTTATGATTTAAAT 473

QY      301 AAAAGGCAGCTGATTACAGAGAGAGATTCMAACAAACACACAGTGGGTCAATGGTCA 360
DB      474 AAAAGGCAGCTGATTACAGAGAGAGATTCMAACAAACACACAGTGGGTCAATGGTCA 533

QY      361 CCAGTGGGTCAATAAATGGCATATGTTTGGAAACAATGACATTTATGTTAAATTTGAACCA 420
DB      534 CCAGTGGGTCAATAAATGGCATATGTTTGGAAACAATGACATTTATGTTAAATTTGAACCA 593

QY      421 AATTATCCAAAGTTACAGAAATCAATGGACGGGAAAGAGATATATATATATATGGAATA 480
DB      594 AATTATCCAAAGTTACAGAAATCAATGGACGGGAAAGAGATATATATATATGGAATA 653

QY      481 ACTGACTGGGTTTATGAAGAGGAAGTCTTCAGTGCCTACTCTGCTCTGTGGTGGTCTCCA 540
DB      654 ACTGACTGGGTTTATGAAGAGGAAGTCTTCAGTGCCTACTCTGCTCTGTGGTGGTCTCCA 713

QY      541 AACGGCACATTTTATGACATATGCCCAATTTTAAACGACAGAGAGTCCCACTTATTGAAATAC 600
DB      714 AACGGCACATTTTATGACATATGCCCAATTTTAAACGACAGAGAGTCCCACTTATTGAAATAC 773

QY      601 TCCTTCTACTCTGATGAGTCACTGACAGTACCCAAAGACTGTAGCGGTTCCATATCCAAAG 660
DB      774 TCCTTCTACTCTGATGAGTCACTGACAGTACCCAAAGACTGTAGCGGTTCCATATCCAAAG 833

QY      661 GAGGAGCTGTGAATCCAACTGTAAAGTTCTTTGTTGTTAAATACAGACTCTCTCAGCTCA 720
DB      834 GCAGGAGCTGTGAATCCAACTGTAAAGTTCTTTGTTGTTAAATACAGACTCTCTCAGCTCA 893

QY      721 GTCAACCAATGCAACTTCCATACAAATCACTGCTCTCTCTTCTATGTTGATAGGGGATCAC 780
```

```
DB      894 GTCAACCAATGCAACTTCCATACAAATCACTGCTCTCTCTTCTATGTTGATAGGGGATCAC 953
QY      781 TACTTGTGTGATGTGACATGGGCAACACAGAAAGAAATTTCTTTGACAGTGGCTCAGGAGG 840
DB      954 TACTTGTGTGATGTGACATGGGCAACACAGAAAGAAATTTCTTTGACAGTGGCTCAGGAGG 1013
QY      841 ATTCAAGAACTATTTCGGTCAATCGATATTGTGACATATGATGAATCCAGTGGAGAGATGGAAC 900
DB      1014 ATTCAAGAACTATTTCGGTCAATCGATATTGTGACATATGATGAATCCAGTGGAGAGATGGAAC 1073
QY      901 TGCTTTAGTGGCACCGGCAACACATTTGAAATAGTACTACTGCGCTGGGTGGAAAGATTTAGG 960
DB      1074 TGCTTTAGTGGCACCGGCAACACATTTGAAATAGTACTACTGCGCTGGGTGGAAAGATTTAGG 1133
QY      961 CTTTCAGAACCTCATTTTACCTCTGATGGTAAATAGTCTTCAACAGATCATCAGCAATGAA 1020
DB      1134 CTTTCAGAACCTCATTTTACCTCTGATGGTAAATAGTCTTCAACAGATCATCAGCAATGAA 1193
QY      1021 GAAGGTTACAGACACATTTTGCTATTTCCTCAAAATAGATAAAAAAGACTGCACATTTATTACA 1080
DB      1194 GAAGGTTACAGACACATTTTGCTATTTCCTCAAAATAGATAAAAAAGACTGCACATTTATTACA 1253
QY      1081 AAAGCACCTCGGGAAGTCAATCGGAGATGAAGCTCTTAAACAGTGAATATCTATACTACATT 1140
DB      1254 AAAGCACCTCGGGAAGTCAATCGGAGATGAAGCTCTTAAACAGCGGATATCTATACTACATT 1313
QY      1141 AGTAATGAATATAAAGGAATGCCAGGAGGAAGGATCTTTATAAAATCCCACTTATTGAC 1200
DB      1314 AGTAATGAATATAAAGGAATGCCAGGAGGAAGGATCTTTATAAAATCCCACTTATTGAC 1373
QY      1201 TATACAAAGTGACATGCTCAGTTGTGAGCTGGAATCCGGAAGGTGTGAGTACTATTCT 1260
DB      1374 TATACAAAGTGACATGCTCAGTTGTGAGCTGGAATCCGGAAGGTGTGAGTACTATTCT 1433
QY      1261 GTGTCATTTCAGTAAAGAGGCGAAGTATTATCAGCTGAGATGTTTCGGTCTCGGCTCGCCC 1320
DB      1434 GTGTCATTTCAGTAAAGAGGCGAAGTATTATCAGCTGAGATGTTTCGGTCTCGGCTCGCCC 1493
QY      1321 CTCTATACTCTACACAGCAGCGTGAATGATTAAGGGCTGAGAGTCTCTGGAAGACAAATTTCA 1380
DB      1494 CTCTATACTCTACACAGCAGCGTGAATGATTAAGGGCGGAGAGTCTCTGGAAGACAAATTTCA 1553
QY      1381 GCTTTGGATATAAATGCTGCAGAAATGTCAGAGTGCCTCCAAAAAACTTGACATTTATTAT 1440
DB      1554 GCTTTGGATATAAATGCTGCAGAAATGTCAGAGTGCCTCCAAAAAACTTGACATTTATTAT 1613
QY      1441 TTGAAATGAAACAAAAATTTTGGTATCAGATGATCTTTGGCTCTCTCATTTTGTATAAATCCAAG 1500
DB      1614 TTGAAATGAAACAAAAATTTTGGTATCAGATGATCTTTGGCTCTCTCCCAATTTTGTATAAATCCAAG 1673
QY      1501 AAATATCTCTACTATTAGATGTTGATGCGGCCCACTGTAGTCAAAAAGCAGACACTGTC 1560
DB      1674 AAATATCTCTACTATTAGATGTTGATGCGGCCCGTGTAGTCAAAAAGCAGACACTGTC 1733
QY      1561 TTCAAGCTGAACCTGGGCCACTTTACCTTGCAGACACAGAAAAATTTATAGTAGTAGCTATTT 1620
DB      1734 TTCAAGCTGAACCTGGGCCACTTTACCTTGCAGACACAGAAAAATTTATAGTAGTAGCTATTT 1793
QY      1621 GATGCGAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGCAATGCAATCAACAGAGACTG 1680
DB      1794 GATGCGAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGCAATGCAATCAACAGAGACTG 1853
QY      1681 GGAACTTTGAAGTTGAAGATCAAAATTAAGCAGCCAGCAATTTTCAAAAATGGGATTT 1740
DB      1854 GGAACTTTGAAGTTGAAGATCAAAATTAAGCAGCCAGCAATTTTCAAAAATGGGATTT 1913
QY      1741 GTGGACAAACAAAGAAATTTGGGCTGGTCAATGAGGGGTACCTAACTCAATG 1800
DB      1914 GTGGACAAACAAAGAAATTTGGGCTGGTCAATGAGGGGTACCTAACTCAATG 1973
QY      1801 GTCTCTGGGATCGGGAAGTGGCGTGTCAAGTGTGAAATAGCCGTGGCGCTGTATCCCCGG 1860
```



[illegible]

Db 132 AGCCGAGAACGTAATCACTAGCTGACTATATTAAGAGATACCTTTCCGGCTCAAGTCTTAC 191  
Qy 61 TCCTTAAGATGATTTTCAGATCATGAATATCTCTACAAACAGAAATAATATCTTGTA 120  
Db 192 TCCTTTGGTGGGTTTCAGACTTTGAATACCTCTACAAACAGAGACAATATCTTGCTG 251  
Qy 121 TTCAATGCTGAATATGGAACAGCTCAGTTTCTTTGGAGAACAGTACATTTGATGAGTTT 180  
Db 252 CTCATGCTGAACATGGAACAGCTCCATTTCTTTGGAGAACAGTACCTTTGGAAGCTTT 311  
Qy 181 GGACATTTCTATCAATGATTTCAATATCTCTGATGGGCAGTTTATTTCTTTAGAATAC 240  
Db 312 GGA-----TATCATTTCACTGTCACCTGACCGACTGTTGTTCTCTCTGGAATAC 359  
Qy 241 AACTAGGTGAACCAATGGAGGCAATTCCTACAGACTTCATATGACATTTATGATTAAT 300  
Db 360 AACTACGTGAAGCAATGGAGACAATTCCTACAGACTTCATACCAATTTATGATGGAAT 419  
Qy 301 AAAAGCAGCTGATTTACAGAGAGAGGATTTCCAAACACACACAGTGGGTACATGGTCA 360  
Db 420 AAAAGCAGCTGATTCAGAGAGAGAGATTTCCAAATATATACAGTGGATACATGGTCA 479  
Qy 361 CCAGTGGGTCTATAAATTTGSCATATGTTTGGAAACAATGACATTTATGTTAAATTTGAACCA 420  
Db 480 CCAGAGGTCTATAAGTTGGCATATGCTCTGGAAGATGATTTTACGTTAAAGTTGAACCA 539  
Qy 421 AATTTACCAAGTTACAGATCATCTGGACGGGGAAGAGATATATATATATATGGAATA 480  
Db 540 CACTTACCTAGTCATAGGATCATCTGACAGGAGAGAAATGTAATATATATATGGAATA 599  
Qy 481 ACTGACTGGGTTTATGAAGAGAGAGTCTTCAGTGGCTACTCTGCTCTGTTGGTGGTCTCCA 540  
Db 600 ACTGACTGGGTTTATGAAGAGAGAGTCTTCGTTGGCTACTCTGCACTGTTGGTGGTCTCCA 659  
Qy 541 AACGGCAGCTTTTATAGCATATGCCAAATTTAAACGACACAGAAAGTCCCACTTATGAATAC 600  
Db 660 AACACACGTTTCTAGCTTATGCCAGTTTAAAGACACAGGAGTCCCGCTCATTTGAATAC 719  
Qy 601 TCCTTCTACTGTGATGATGCTACGCGATACCAAGACATGTAAGGGTTCCATATCCAAAG 660  
Db 720 TCCTTCTATTTCTGATGATGCTACGTCAGTACCCCAAGACAGTGTGGAATCCATACCCCAAG 779  
Qy 661 CGAGAGCTGTGAATCCAACTGTAAGTTCTTTTGTAAATPACAGACTCTCTCAGCTCA 720  
Db 780 CGAGAGCTGTGAATCCAACTGTAAGTTCTTTTGTAAATPACAGACTCTCTCAGCTCA 839  
Qy 721 GTCACCAATGCAACTTCCATACAAATCACTGCTCTGCTTCTATGTTGATAGGGGATCAC 780  
Db 840 TCCTCTAGTGGCTCCCATCCAAATCCCTGCTCCTGCACTCTGTGCAAGAGGGGATCAC 899  
Qy 781 TACTTGTGTGATGTCATGGGCAACAGAAAGAAATTTCTTTTGGAGTGGCTCAGGAGG 840  
Db 900 TATTTATGTGATGTGTGGGCTACAGAAAGAAATTTTCACTACAGTGGCTCAGGAGG 959  
Qy 841 ATTCAGACTATTCGGTCATGNTATTTGTGACTATGATGATGATCCAGTGGAGATGGAC 900  
Db 960 ATTCAGACTATTCGGTGTAGGCTATCTGTGACTATGATGATGATCAACCTAACGTTGAAC 1019  
Qy 901 TCCTTAGTGGCAGGCAACATTTGAAATGAGTACTACTGGCTGGGTTTGGAAAGATTTAGG 960  
Db 1020 TGTCCATCCGAGGAGCAGCATGTTGNAATGATGATACCAAGGCTGGGTCCGAGAGATTTAGG 1079  
Qy 961 CTTTCAGAACTCTCATTTTACCCTTGATGTAATAGCTTCTCAAGATCATACGAAATGAA 1020  
Db 1080 CCGCAGAACTCTCATTTCACTCTGATGGAAGCAGCTTCTATAAGATCATACGCGACAA 1139  
Qy 1021 GAAGGTTACAGACATTTGCTATTTTCCAAATAGATATAAAGACTGCAATTTATTACA 1080  
Db 1140 GATGGCTACAAACATCTGCCATTTCCCGAAGATAAGAAAGCTGTACATTTATTACA 1199  
Qy 1081 AAAGGCCTCGGAGCTCATCGGATPAGAAGCTCTAACCAAGTATTTACTACTACATTT 1140  
Db 1200 AAAGGACCTGGGAGTCTATAGTATCGAAGCTCTGACCGAGATTTATCTACTACTACATTT 1259

Qy 1141 AGTAATGAATATAAAGGAATGCCAGGAGGAAGGAATCTTTATAAAATCCAACTTATTGAC 1200  
Db 1260 AGTAACCAATATAAAGAAATGCCAGGAGGAAGAAATCTCTATAAAATTCNACTTACTGAC 1319  
Qy 1201 TATACAAAGTGACATGCTTCAGTTGTGAGCTGAATCCGGAAGGTTGTCGATCTATTCT 1260  
Db 1320 CACACAAATGTGAAGTGCCTTAGTTGTGACCTGAATCCAGAAAGATGTCAATATTATGCG 1379  
Qy 1261 GTGTCACTTCAGTAAGAGCGAAGTATTTATCAGCTGAGATGTTCCGGTCTTGGTCTGCCC 1320  
Db 1380 GTATCATTTAGTAAGAGGCAAGTACTATCAGCTGGGATGTTGGGGCCCCGGGTCTGCCC 1439  
Qy 1321 CTCTATATCTCTACACAGCAGCGTGAATGAATAAAGGCTGAGAGTCTCTGGAAGACAATCA 1380  
Db 1440 CTCTACATCTCTACATCTGTAACGATCATTAAGAGCTGCGAGTCTCTGGAAGACAATCT 1499  
Qy 1381 GCTTTGGATAAAATGCTGCAGAAATGTCAGATGCCCTCCAAAACCTGGACTTCAATTT 1440  
Db 1500 GCTTTGGATAGATGCTGCAGGATGTCAGATGCCCTTCAAAAAAATTTGACCTTCAATGTT 1559  
Qy 1441 TTGAATGAAACAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGTATATAATCCAG 1500  
Db 1560 TTGAATGAAACAAAGATTTTGGTATCAAAATGATCTTGGCCCCCTCATTTTGTATAAATCCAAG 1619  
Qy 1501 AAATATCTCTCTACTATTATAGATGTGTATGACGCCCATGTAGTCAAAAAAGCAGACACTGTC 1560  
Db 1620 AAATATCTCTCTACTATTATAGATGTATGCAAGTCCCTGTAGTCAAAAAAGCAGACTGCTCC 1679  
Qy 1561 TTCAGACTGAACTGGGCCACTTACCTTGCAAGCAGAAAAACATTTATATAGTACTAGCTTT 1620  
Db 1680 TTCAGACTGAACTGGGCCACTTACCTTGCAAGTACAGAAAAACATCATATAGTACTAGCTTT 1739  
Qy 1621 GATGCGAGAGAGTGGTTACCAAGGAGATAAGATCATGCAATCAACAGAGACTG 1680  
Db 1740 GACGCGCAGAGAGTGGTTTACCAGGAGATAAGATCATGCAATCAACAGAGAGATTTG 1799  
Qy 1681 GGACATTTGAAGTTGAAGATCAAAATTCGAAGCAGCAGACAAATTTTCAAAAAATGGATTT 1740  
Db 1800 GGAACTCTGGAGTTGAAGATCAAAATTCGAAGCAGCAGGCAATTTTGTAAAAATGGGATTT 1859  
Qy 1741 GTGGACAAACAAAGAAATTTGGGGCTGGTCAATATGGAAGGTAAGTAACTCAATG 1800  
Db 1860 GTGGATACCAAGCGAGTTTGAATTTGGGGCTGGTCAATATGGAAGGTAAGTAACTCAATG 1919  
Qy 1801 GTCTGGGATCGGGAAGTGGGCTGTTCAAGTGTGGAATAGCGCTGGCGCTGTATCCCGG 1860  
Db 1920 GTCTGGGATCGGGAAGTGGGCTGTTCAAGTGTGGAATAGCTGTGGCACCTGTGTACCGG 1979  
Qy 1861 TGGGAGTACTATGACTCAGTGTACACAGAACTTTACATGGGTCTCCCACTCCAGAGAC 1920  
Db 1980 TGGGAGTACTATGACTCAGTGTACAGAGGTTACATGGGTCTCCCAATTCAGAGAC 2039  
Qy 1921 AACCTTGACCAATTAAGAAAAATTCACAGTCAATGAGCAGAGCTGAAAAATTTTAAACAGTT 1980  
Db 2040 AACCTTGACCAATTAAGAAAAATTCACAGTCAATGAGCAGAGCTGAAAAATTTTAAACAGTT 2099  
Qy 1981 GAGTACCTCTCTTATTCATGGAACAGCAGATGATAAGCTTCACTTTTCCAGAGTCAAGTCA 2040  
Db 2100 GAGTACCTCTCTTATTCATGGAACAGCAGATGATAAGCTTCACTTTTCCAGAGTCAAGTCA 2159  
Qy 2041 ATCTCCAAAGCCCTGGTCTGATGTTGGAGTTCAGGCAATTCAGGCAATGTTATCTGATGAA 2100  
Db 2160 ATCTCCAAAGCCCTGGTCTGATGTTGGAGTTCAGGCAATGTTGTTACAGGATGAA 2219  
Qy 2101 GACCAATGAATAGCTAGCAGCAGCAGCACAACCAATATATATACCAATGAGCCACTTC 2160  
Db 2220 GACCAATGAATAGCTAGCAGCAGCAGCACAACCAATATATATACCAATGAGCCACTTC 2279  
Qy 2161 ATAAACCAATGTTTCTCTTTAC 2182  
Db 2280 CTCAGCAGCTGCTTCTCTTAC 2301

```
RESULT 3
AY418368
LOCUS
DEFINITION Homo sapiens DPP4 gene, VIRTUAL TRANSCRIPT, partial sequence,
                2193 bp DNA linear GSS 17-DEC-2003
                genomic survey sequence.
ACCESSION AY418368
VERSION AY418368.1 GI:39774328
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
            1 (bases 1 to 2193)
            Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
            Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
            Science 302 (5652), 1960-1963 (2003)
            2 (bases 1 to 2193)
            Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
            Direct Submission
            Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
            These sequences were made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES
    source      1..2193
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                <1..>2193
                /gene="DPP4"
                /locus_tag="HCM6534"
ORIGIN
Query Match      70.9%; Score 1549.4; DB 10; Length 2193;
Best Local Similarity 77.3%; Pred. No. 0;
Matches 1550; Conservative 0; Mismatches 456; Indels 0; Gaps 0;
Qy 80 ATCATGATATCTCTACAAACAGAAATAATATCTTGGTATTCATGCTGAATATGGAA 139
Db 188 ATCATGATATCTCTACAAACAGAAATAATATCTTGGTATTCATGCTGAATATGGAA 247
Qy 140 ACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTGGACATTCATCAATGATT 199
Db 248 ACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTGGACATTCATCAATGATT 307
Qy 200 ATTCATATCTCTGATGGGAGTTTATTTCTCTTAGAATACAACTACGTGAAGCAATGGA 259
Db 308 ATTCATATCTCTGATGGGAGTTTATTTCTCTTAGAATACAACTACGTGAAGCAATGGA 367
Qy 260 GGCATTCTACACAGCTTCATATGACATTTATGATTTAAATAAAGGCGAGCTGATTACAG 319
Db 368 GGCATTCTACACAGCTTCATATGACATTTATGATTTAAATAAAGGCGAGCTGATTACAG 427
Qy 320 AAGAGAGATTCCAAACACACAGTGGGTGCATGTCACATGTCACCGTGGTCAATAATGG 379
Db 428 AAGAGAGATTCCAAACACACAGTGGGTGCATGTCACATGTCACCGTGGTCAATAATGN 487
Qy 380 CATATGTTTGGAAACAAAGACATTTATGTTTAAATTTGAACCAAAATTTACCAAGTTACAGAA 439
Db 488 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 547
Qy 440 TCACATGACGGGGAAGAAGATATATATATATGGAATAACTGACTGGGTTTATGAAG 499
548 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 607
500 AGGAAGTCTTCAGTGCCTACTCTGCTCTGTGGTGTCTCCAAACGGCAGCTTTTGTAGCAT 559
608 AGGAAGTCTTCAGTGCCTACTCTGCTCTGTGGTGTCTCCAAACGGCAGCTTTTGTAGCAT 667
560 ATGCCCAATTTAAACGACACAGAGTCCCACTTATTGAATATCTCTTCTACTCTGATGAGT 619
668 ATGCCCAATTTAAACGACACAGAGTCCCACTTATTGAATATCTCTTCTACTCTGATGAGT 727
620 CACTGCGATACCCAAAGAGTCTGAGGGTTCATATCCAAAGCGCAGAGCTGTGAATCCAA 679
728 CACTGCGATACCCAAAGAGTCTGAGGGTTCATATCCAAAGCGCAGAGCTGTGAATCCAA 787
680 CTGTAAGTCTTCTTGTGTAATAACAGACTCTCTCAGCTCAGTCAACCAATGCAACTTCCA 739
788 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 847
740 TACAAATCAGTCTCTGCTTCTATGTTGATAGGGGATCATCTACTTGTGTGATGTGACAT 799
848 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 907
800 GGGCAACACAAAGAAATTTCTTTCAGTGCCTCAGGAGATTCAGAACTATTCGGTCA 859
908 GGGCAACACAAAGAAATTTCTTTCAGTGCCTCAGGAGATTCAGAACTATTCGGTCA 967
860 TGGATATTTGTGACTATGATCAATCCAGTGAAGATGAACTGCTTAGTGGCAGCGCAAC 919
968 TGGATATTTGTGACTATGATCAATCCAGTGAAGATGAACTGCTTAGTGGCAGCGCAAC 1027
920 ACATGAAATGAGTACTACTGCTGGGTGGAGATTTAGGCTCTCAGAACCTCATTTTCA 979
1028 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1087
980 CCCTTGATGGTAAATAGCTTCTACAAGATCATCAGCAATGAAGAGGTTACAGACATTT 1039
1088 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1147
1040 GCTATTTCCAAATAGATAAAAGAGCTGCACATTTATTACAAAGGCACTCGGGAAGTCA 1099
1148 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1207
1100 TCGGGATAGAGCTCTAACAGTGATTTATCTACTACTAGTAAATGAATATAAAGGAA 1159
1208 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1267
1160 TGCAGAGGAGGAACTTTTATAAATCCAACTATTGACTATACAAAGTGAATGATGCC 1219
1268 TGCAGAGGAGGAACTTTTATAAATCCAACTATTGACTATACAAAGTGAATGATGCC 1327
1220 TCAGTTGTGAGCTCAATCCGGAAGGTGTGAGTACTATTCTGTGTCATTTCAGTAAAGAGG 1279
1328 TCAGTTGTGAGCTCAATCCGGAAGGTGTGAGTACTATTCTGTGTCATTTCAGTAAAGAGG 1387
1280 CGAAGTATTATCAGCTGAGATGTTCCGGTCTCGCTCTGCCCTCTATACTCTTACACAGCA 1339
1388 CGAAGTATTATCAGCTGAGATGTTCCGGTCTCGCTCTGCCCTCTATACTCTTACACAGCA 1447
1340 GCGTGAATGATAAAGGCTGAGAGTCTGGAAGCAATTCAGCTTTGGATATAAATGCTGCG 1399
1448 GCGTGAATGATAAAGGCTGAGAGTCTGGAAGCAATTCAGCTTTGGATATAAATGCTGCG 1507
1400 AGAATGTCAGATGCCCTCCAAAGAACTGGAGTTCATTATTTCGAATGAACAAATTTT 1459
1508 AGAATGTCAGATGCCCTCCAAAGAACTGGAGTTCATTATTTCGAATGAACAAATTTT 1567
1460 GGTATCAGATGATCTTGCCTCTCATTTTGAATAATCCAGAAATATCTCTTACTATTAG 1519
1568 GGTATCAGATGATCTTGCCTCTCATTTTGAATAATCCAGAAATATCTCTTACTATTAG 1627
1520 ATGTGATGCGGGCCCATGTAGTCAAAAGCAGACATGCTTTCAGACTGAATCGGGCCA 1579
1628 ATGTGATGCGGGCCCATGTAGTCAAAAGCAGACATGCTTTCAGACTGAATCGGGCCA 1687
```



[illegible]









1454 ATTTTGGTATCAGATGATCTTGGCTCCTCATTTTGTATTAATCCAGAAATATCTCTAC 1513  
 Db ATTTTGGTATCAGATGATCTTGGCTCCTCATTTTGTATTAATCCAGAAATATCTCTAC 748  
 1514 TATTAGATGTGTATGAGGCCCATGTAGTCAAAAAGCAGACACTGTCTTCAGACTGAACT 1573  
 Db TATTAGATGTGTATGAGGCCCATGTAGTCAAAAAGCAGAACTGTCTTCAGACTGAACT 688  
 1574 GGGCCACTTACCTTCAAGCAGACAGAAAACATTATAGTACTAGCTTTGTATGTCAGAGGAA 1633  
 Db GGGCCACTTACCTTCAAGCAGACAGAAAACATTATAGTACTAGCTTTGTATGTCAGAGGAA 628  
 1634 GTGGTTACCAAGGATAGATCATGTCATGCAATCAACAGAGACTGGGAACATTTGAG 1693  
 Db GTGGTTACCAAGGATAGATCATGTCATGCAATCAACAGAGACTGGGAACATTTGAG 568  
 1694 TTGAAGATCAAAATGAAGCAGCAGACAAATTTTCAAAAATGGATTTTGGGACAAAC 1753  
 Db TTGAAGATCAAAATGAAGCAGCAGACAAATTTTCAAAAATGGATTTTGGGACGACAAAC 508  
 1754 GAATTCGAATTTGGGGCTGTCATATGAGGGTACGTAACCTCAATGCTTGGGATCGG 1813  
 Db GAATTCGAATTTGGGGCTGTCATATGAGGGTACGTAACCTCAATGCTTGGGATCGG 448  
 1814 GAAGTGGCGGTTCAGTGTGGAATAGCCGTGGCGCTGTATCCGGTGGGAGTACTATG 1873  
 Db GAAGTGGCGGTTCAGTGTGGAATAGCCGTGGCGCTGTATCCGGTGGGAGTACTATG 388  
 1874 ACTCAGTGTACACAGAACTTACATGGGTCTCCCACTCCAGAGACAACTTGACCAAT 1933  
 Db ACTCAGTGTACACAGAACTTACATGGGTCTCCCACTCCAGAGACAACTTGACCAAT 328  
 1934 ACAGAAATTCAGACAGCTCATGAGCAGCTGAATAATTTTAAACAGAGTGTAGTACCTCTTA 1993  
 Db ACAGAAATTCAGACAGCTCATGAGCAGCTGAATAATTTTAAACAGAGTGTAGTACCTCTTA 268  
 1994 TTCATGGAACACAGATGATACCTTACCTTTCAGCAGCTCAGATCTCAAAAGCC 2053  
 Db TTCATGGAACACAGATGATACCTTACCTTTCAGCAGCTCAGATCTCAAAAGCC 208  
 2054 TGGTCGATGTCGAGTGGATTTCCAGGCAATGTTGATATGATGAAAGCAGCTGAATAG 2113  
 Db TGGTCGATGTCGAGTGGATTTCCAGGCAATGTTGATATGATGAAAGCAGCTGAATAG 148  
 2114 CTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2173  
 Db CTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 88  
 2174 TCTCTTTACCT 2184  
 Db TCTCTTTACCT 77

RESULT 8  
 LOCUS BX348426 886 bp mRNA linear EST 08-APR-2004  
 DEFINITION BX348426 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 clone CS0DI084YJ15 5-PRIME, mRNA sequence.  
 ACCESSION BX348426  
 VERSION BX348426.1 GI:30375280  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homidae; Homo.  
 REFERENCE 1 (bases 1 to 886)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: seq@genoscope.cns.fr, Web: www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with NotI and cloned  
 into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 1082.r  
 For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?c=CS0BAP032ZD12\_AF03069\_2&c=1082.r

FEATURES  
 source  
 Location/Qualifiers  
 1. 886  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="caxon:9606"  
 /clone="CS0DI084YJ15"  
 /issue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with NotI and cloned into the NotI and EcoRV  
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
 Query Match 38.0%; Score 829.2; DB 5; Length 886;  
 Best Local Similarity 98.2%; Pred. No. 4.3e-208;  
 Matches 859; Conservative 0; Mismatches 14; Indels 2; Gaps 2;  
 QY 1260 TGTGTCTTCTAGTAAAGAGGCGAAGTATTATCAGCTGAGATGTTCCGGTCTGGTCTGCC 1319  
 Db 1 TGTGTCTTCTAGTAAAGAGGCGAAGTATTATCAGCTGAGATGTTCCGGTCTGGTCTGCC 60  
 QY 1320 CCTCTATCTCTACACAGCAGCGTGAATGATAAAGGGCTGAGAGTCTCTGGAAGACAAATTC 1379  
 Db 61 CCTCT-TACTCTACACAGCAGCGTGAATGATAAAGGGCTGAGAGTCTCTGGAAGACAAATTC 119  
 QY 1380 AGCTTTGATGATAAATGCTGCAAGATGTCAGATGCTCCCAAAAACCTGACCTTCAATAT 1439  
 Db 120 AGCTTTGATGATAAATGCTGCAAGATGTCAGATGCTCCCAAAAACCTGACCTTCAATAT 179  
 QY 1440 TTTGAATGAAAACAAAATTTTGGTATCAGATGATCTTGCTCTCTCATTTTGTATAAATCCAA 1499  
 Db 180 TTTGAATGAAAACAAAATTTTGGTATCAGATGATCTTGCTCTCTCATTTTGTATAAATCCAA 239  
 QY 1500 GAAATATCTCTCTATATTAGATGTTGAGGCCCATGTAGTCAAAAAGCAGACACTGT 1559  
 Db 240 GAAATATCTCTCTATATTAGATGTTGAGGCCCATGTAGTCAAAAAGCAGACACTGT 299  
 QY 1560 CTTTCAGACTGAACCTGGGCCACTTACCTTGCAGCAGACAGAAAACATTTATAGTAGCTAGCTT 1619  
 Db 300 CTTTCAGACTGAACCTGGGCCACTTACCTTGCAGCAGCAGAAAACATTTATAGTAGCTAGCTT 359  
 QY 1620 TGATGGCAGAGGAGTGGTTTACCAAGGAGATAAGATCATGTCATCAACAGAGAGACT 1679  
 Db 360 TGATGGCAGAGGAGTGGTTTACCAAGGAGATAAGATCATGTCATCAACAGAGAGACT 419  
 QY 1680 GGGAACTTTGAGTTGAAGATCAAAATGAAGCAGCAGACAAATTTTCAAAAATGGGAT 1739  
 Db 420 GGGAACTTTGAGTTGAAGATCAAAATGAAGCAGCAGACAAATTTTCAAAAATGGGAT 479  
 QY 1740 TGTGGACACAAACCAATTTGGGGCTGGTTCATATGAGGGTACGTAACCTTCAAT 1799  
 Db 480 TGTGGACACAAACCAATTTGGGGCTGGTTCATATGAGGGTACGTAACCTTCAAT 539  
 QY 1800 GGTCTCTGGGATCGGGAAGTGGCGTGTTCAGAGTGTGGAATAGCCGTGGCGCTGTATCCCG 1859  
 Db 540 GGTCTCTGGGATCGGGAAGTGGCGTGTTCAGAGTGTGGAATAGCCGTGGCGCTGTATCCCG 599  
 QY 1860 GTGGGAGTATATGATCTCAGTGTACAGAGAGCTTACATGGGTCTCCGAACTCCAGAGAA 1919  
 Db 600 GTGGGAGTATATGATCTCAGTGTACAGAGAGCTTACATGGGTCTCCGAACTCCAGAGAA 659

Qy	1920	CAACCTTGACCATTTACAGAAATTTCAACAGTCATGACGACAGCTGAAATTTTAAACAAGT	1979
Db	660	CAACCTTTGACCATTTACAGAAATTTCAAACAGTCATGACGACAGCTGAAATTTTAAACAAGT	719
Qy	1980	TGAGTACCTCCTTTATTCATGGAACAGCAGAGATGATAGTTTCATCTTTCAGCAGTCAGCTCA	2039
Db	720	TGAGTACCTTCTTTATTCATGGAACAGCAGAGATGATAGTTTCATCTTTCAGCAGTCAGCTCA	779
Qy	2040	GATCTCCAAAGCCCTGGTTCGATGTGGAGTGGGATTTCCAGGCCAATGGGTATCTGATGA	2099
Db	780	GATCTCCAGCCCTGGTTCGATGTGGAGTGGGATTTCCAGGCCAATGGGTATCTGATGA	839
Qy	2100	AGACCATGGAAATAGCTAGCAGCAGCAGCACACCAAC	2134
Db	840	AGACCATGG-ATAGCTAGCAGCAGCAGCACACCAAC	873

RESULT 9	AL548644	1014 bp	mrna	linear	EST 25-MAR-2004
LOCUS	AL548644				
DEFINITION	AL548644 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0DI036YE04 5-PRIME, mRNA sequence.				
ACCESSION	AL548644				
VERSION	AL548644.3	GI:45749064			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1. (bases 1 to 1014)				
AUTHORS	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished (2001)				
COMMENT	On Feb 15, 2001 this sequence version replaced gi:31270470.				

```

FEATURES
source
1. .1014
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1036YE04"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN	Query Match	Best Local Similarity	Matches	Conservative	Score	DB 1;	Length	Indels	Gaps
		95.7%;	Pred. No. 4.2e-205;	19; Mismatches	14;	Indels	6;	Gaps	5;
Qy	1	AGTGGCAAACTTCACATCCTAACGTGATTACTTTAAAAAATACTTATAGACTGAAGTTATAC	60						
Dd	111	AGTGGCAAACTTCACATCCTAACGTGATTACTTTAAAAAATACTTATAGACTGAAGTTATAC	170						
Qy	61	TCCTTAAGATGGATTTCAGATCATGAATATCTCTACAACAAGAATAAATATCTTGTTA	120						
Dd	171	TCCTTAAGATGGATTTCAGATCATGAATATCTCTACAACAAGAATAAATATCTTGTTA	230						

Qy	121	TTCAATGCTGAAATATGGAAACAGCTCAGTTTTCTTGGAGAACAGTACATTTGATGAGTTT	180
Db	231	TTCAATGCTGAAATATGGAAACAGCTCAGTTTTCTTGGAGAACAGTACATTTGATGAGTTT	290
Qy	181	GGACATTTCTATCAATGATATTCAATATCTCCTGATGGGAGTTTATCTCTTAGAATAC	240
Db	291	GGACATTTCTATCAATGATATTCAATATCTCCTGATGGGAGTTTATCTCTTAGAATAC	350
Qy	241	AACTACGTGAAGCAATGGAGGCAATTCCTACACAGCTTTCATATGACATTTATGATTTAAAT	300
Db	351	AACTACGTGAAGCAATGGAGGCAATTCCTACACAGCTTTCATATGACATTTATGATTTAAAT	410
Qy	301	AAAAGCAGCTGATTACAGNAGAGAGATTCGAAACAAACACACAGTGGGTCCATGCTCA	360
Db	411	AAAAGCAGCTGATTACAGNAGAGAGATTCGAAACAAACACACAGTGGGTCCATGCTCA	470
Qy	361	CCAGTGGGTTCATAAAATGGCATATGTTTGGAAACAATGACATTTATGTTAAAAATTGAACCA	420
Db	471	CCAGTGGGTTCATAAAATGGCATATGTTTGGAAACAATGACATTTATGTTAAAAATTGAACCA	530
Qy	421	AATTTACCAAGTTACAGAAATCACATGGACGGGGAAAGAGATATAATATATAATGGAATA	480
Db	531	AATTTACCAAGTTACAGAAATCACATGGACGGGGAAAGAGATATAATATATAATGGAATA	590
Qy	481	ACTGACTGGGTTTATGAAGGGAAGTCTTCAGTGCCTCTACTCTGCTGTGGTGGTCTCCA	540
Db	591	ACTGACTGGGTTTATGAAGGGAAGTCTTCAGTGCCTCTACTCTGCTGTGGTGGTCTCCA	650
Qy	541	AACGGCACATTTTTTACCATATGCCCAATTTTAAACGACACAGAAGTCCCACTTATTTGAATAC	600
Db	651	AACGGCACATTTTTTACCATATGCCCAATTTTAAACGACACAGAAGTCCCACTTATTTGAATAC	710
Qy	601	TCCTTCTACTGTATGAGTCACTGCAAGTACCCAAAGACTGTACGGGTTCCATATCCAAAG	660
Db	711	TCCTTCTACTGTATGAGTCACTGCAAGTACCCAAAGACTGTACGGGTTCCATATCCAAAG	770
Qy	661	GCAGGAGCTGTGAATCCCAACTGTAAAGTCTTTTGTGTAAATACAGACTCTCTCAGCTCA	720
Db	771	GCAGGAGCTGTGAATCCCAACTGTAAAGTCTTTTGTGTAAATACAGACTCTCTCAGCTCA	830
Qy	721	GTCAACCAATGCCAATTTCCCATACAAATCACTGCTCTCTGCTTCTATGTTGATAGGGATCAC	780
Db	831	GTCAACCAATGCCAATTTCCCATACAAATCACTGCTCTCTGCTTCTATGTTGATAGGGAT-MC	889
Qy	781	TACTTGTGTGATGTGACATGGGGCAACAAAGAAAGAAATTTCTTGTGAGTGGCTCAGGAGG	840
Db	890	TACTTGTGTGATGTGACATGGGGCAACAAAGAAAGAAATTTCTTGTGAGTGGCTCAGGAGG	947
Qy	841	ATTTCAGAACTATTTCGGTTCATGGATATTTGTGACTATGATGAATCCAGTGGAGATCGAAC	900
Db	948	ATACRG-ACATATTCGGTCAATGG-TATTTCGAGTCTGCTGTTT-CHGWRGGARATGGGAM	1004
Qy	901	TGCTTAGTGG 910	
Db	1005	TRMTTAKTGG 1014	

RESULT 10	AL5411949	1022 bp	mRNA	linear	EST 24-MAR-2004
LOCUS	AL5411949				
DEFINITION	AL5411949 Homo sapiens PLACENTA Homo sapiens CDNA clone CS0DE007YN21 5-PRIME, mRNA sequence.				
ACCESSION	AL5411949				
VERSION	AL5411949.3	GI:45717525			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 1022)				
AUTHORS	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.				
TITLE	Full-length CDNA libraries and normalization				

JOURNAL  
COMMENT

Unpublished (2001)  
On Feb 15, 2001 this sequence version replaced gi:30546617.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
was not normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
This sequence belongs to sequence cluster 1082.r  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?s=CS0DE007CG11QPl6c=1082.r.

FEATURES  
source

1..1022  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DE007YN21"  
/tissue\_type="PLACENTA"  
/clone\_lib="Homo sapiens PLACENTA"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo(dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."

ORIGIN

Query Match 36.8%; Score 803.2; DB 1; Length 1022;  
Best Local Similarity 94.5%; Pred. No. 3.5e-201;  
Matches 843; Conservative 11; Mismatches 33; Indels 5; Gaps 2;

QY 1 AGTCGCAAACTTACACTCTAATGATTCTTAAATAAATCTTATAGACTGAAGTTATAC 60  
DB 126 AGTCGCAAACTTACACTCTAATGATTCTTAAATAAATCTTATAGACTGAAGTTATAC 185  
QY 61 TCCTTAAGATGATTTTCAGATCATGATATCTCTACAAACAGAAATATATCTTGGTA 120  
DB 186 TCCTTAAGATGATTTTCAGATCATGATATCTCTACAAACAGAAATATATCTTGGTA 245  
QY 121 TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGAAACAGTACATTTGATGAGTTT 180  
DB 246 TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGAAACAGTACATTTGATGAGTTT 305  
QY 181 GGACATTTCTCAATGATTATTCATATCTCTGATGGGCGATTTATCTCTTAGAATAC 240  
DB 306 GGACATTTCTCAATGATTATTCATATCTCTGATGGGCGATTTATCTCTTAGAATAC 365  
QY 241 AACTAGCTGAAGCAATGGAGGCAATTCCTACACAGCTTCATATGACATTTATGATTTAAAT 300  
DB 366 AACTAGCTGAAGCAATGGAGGCAATTCCTACACAGCTTCATATGACATTTATGATTTAAAT 425  
QY 301 AAAAGGCAGCTGATTACAGAGAGAGGATTCCTACAAACAGACAGGCGGTCACTGGTCA 360  
DB 426 AAAAGGCAGCTGATTACAGAGAGAGGATTCCTACAAACAGACAGGCGGTCACTGGTCA 485  
QY 361 CCAGTGGGTCATAAAATGGCATATGTTTGGAAACAAATGACATTTATGTTAAATTTGAACCA 420  
DB 486 CCAGTGGGTCATAAAATGGCATATGTTTGGAAACAAATGACATTTATGTTAAATTTGAACCA 545  
QY 421 AATTTACCAAGTTTACAGATTCATGAGCGGGGAAAGAGATATATATATATATGGAATTA 480  
DB 546 AATTTACCAAGTTTACAGATTCATGAGCGGGGAAAGAGATATATATATATATGGAATTA 605  
QY 481 ACTGACTGGGTTTATCAAGAGGAGCTTCAGTGCCTACTCTGCTCTGCTGCTGCTCCCA 540  
DB 606 ACTGACTGGGTTTATCAAGAGGAGCTTCAGTGCCTACTCTGCTCTGCTGCTGCTCCCA 665  
QY 541 AACGGCAGCTTTTATAGCATATGCCCAATTTTAAACGACACAGAGTCCCACTTTATGGAATAC 600  
DB 666 AACGGCAGCTTTTATAGCATATGCCCAATTTTAAACGACACAGAGTCCCACTTTATGGAATAC 725

QY 601 TCCTTCTACTCTGATGAGTCACTGAGTACCCAAAGACTGTACGGGTTCCATATCCAAAG 660  
DB 726 TCCTTCTACTCTGATGAGTCACTGAGTACCCAAAGACTGTACGGGTTCCATATCCAAAG 785  
QY 661 GCAGGAGCTGTGAATCCAACTGTAAGTTCTTTGTTGTAATACAGACTCTCTCAGCTCA 720  
DB 786 GCAGGAGCTGTGAATCCAACTGTAAGTTCTTTGTTGTAATACAGACTCTCTCAGCTCA 845  
QY 721 GTCACCAATGCAACTTCCATCAAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780  
DB 846 GTCACCAATGCAACTTCCATCAAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 905  
QY 781 TACTTGTGTGATGTGACATGCGCAACACAAAGAAATTTCTTTCAGTGGTCTCAGGAG 840  
DB 906 TACTTGTGTGATGTGACATGCGG--CAMACAAGAAATTTCTTTCAGTGGTCTCAGGAG 962  
QY 841 ATTACAGAACTATTTCGCTCATGGATATTGTGTGATGATGATGATGATGATGATGATG 892  
DB 963 --TTCAGACTATTTCGCTCATGGATATTGTGTGATGATGATGATGATGATGATGATG 1012

RESULT 11  
AL544456  
LOCUS  
DEFINITION  
AL544456 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
clone CS0I018YJ07 5-PRIME, mRNA sequence.  
ACCESSION  
VERSION  
AL544456.3 GI:45744959  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
On Feb 15, 2001 this sequence version replaced gi:31266300.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
1082.r  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?s=CS0DI018CE04QP16c=1082.r.

FEATURES  
source

1..1007  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DI018YJ07"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
1082.r

ORIGIN

Query Match 36.4%; Score 795.8; DB 1; Length 1007;  
Best Local Similarity 95.4%; Pred. No. 3.2e-199;  
Matches 852; Conservative 9; Mismatches 26; Indels 6; Gaps 4;

QY 1 AGTCGCAAACTTACACTCTAATGATTCTTAAATAAATCTTATAGACTGAAGTTATAC 60  
DB 114 AGTCGCAAACTTACACTCTAATGATTCTTAAATAAATCTTATAGACTGAAGTTATAC 173

```
QY 61 TCCTTAAGATGATTTTCAGATCATGATATCTCTACAAACAGAAATAATATCTTGTTA 120
Db 174 TCCTTAAGATGATTTTCAGATCATGATATCTCTACAAACAGAAATAATATCTTGTTA 233
QY 121 TTCAATGCTGAATATGGAACAGCTCAGTTTTCTTGGAGAACAGTACATTTTATGATGATTT 180
Db 234 TTCAATGCTGAATATGGAACAGCTCAGTTTTCTTGGAGAACAGTACATTTTATGATGATTT 293
QY 181 GGACATTTCTATCAATGATTTATTTCAATATCTCTGATGGCAGATTTTATTTCTTTAGATATAC 240
Db 294 GGACATTTCTATCAATGATTTATTTCAATATCTCTGATGGCAGATTTTATTTCTTTAGATATAC 353
QY 241 AACTACGTGAACATGAGGATTTCTTACAGCTTCTATGATGATGATTTATGATTTTAAAT 300
Db 354 AACTACGTGAACATGAGGATTTCTTACAGCTTCTATGATGATGATTTTATGATTTTAAAT 413
QY 301 AAAAGGCGAGCTGATTTACAGAGAGAGAGGATTTCCAAACACACACAGTGGGTACATGTCAC 360
Db 414 AAAAGGCGAGCTGATTTACAGAGAGAGGATTTCCAAACACACACAGTGGGTACATGTCAC 473
QY 361 CCAGTGGGTCTATAATTTGGCATATGTTTGGAAACATGACATTTATGTTTAAATTTGAACCA 420
Db 474 CCAGTGGGTCTATAATTTGGCATATGTTTGGAAACATGACATTTATGTTTAAATTTGAACCA 533
QY 421 AATTATCCAGTTACAGATTCATGAGCGGGAAGAGATATATATATATATATATATATATATAT 480
Db 534 AATTATCCAGTTACAGATTCATGAGCGGGAAGAGATATATATATATATATATATATATATAT 593
QY 481 ACTGACTGGGTTTATGAGAGGAGTCTTCAGTGGCTACTCTGCTCTGCTGGTGGTCTTCCA 540
Db 594 ACTGACTGGGTTTATGAGAGGAGTCTTCAGTGGCTACTCTGCTCTGCTGGTGGTCTTCCA 653
QY 541 AACGGCCTTTTTTAGCATATGCCCCAATTTAAACACACAGAGTCCCACTTTATTTGAATATAC 600
Db 654 AACGGCCTTTTTTAGCATATGCCCCAATTTAAACACACAGAGTCCCACTTTATTTGAATATAC 713
QY 601 TCCTTCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 714 TCCTTCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 773
QY 661 GCAGAGCTGTAATCCCACTGTAAGTTCTTTTGTGT -AAATACAGACTCTCTCAGCTC 719
Db 774 GCAGAGCTGTAATCCCACTGTAAGTTCTTTTGTGT -AAATACAGACTCTCTCAGCTC 833
QY 720 AGTCACCAATGCACTTCCATACAAATCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 779
Db 834 AGTCACCAATGCACTTCCATACAAATCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 893
QY 780 CTACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 839
Db 894 CTACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 950
QY 840 GATTCAGACTATTTGGTCAATGATATTTTGGATGATGATGATGATGATGATGATGATGATGAT 892
Db 951 G- -TTCAGATTAATTCGGTCTGCTGATTTTGTGATGATGATGATGATGATGATGATGATGAT 1001
```

```
RESULT 12
LOCUS AUI22649 846 bp mRNA linear EST 01-AUG-2002
DEFINITION AUI22649 MAMMA1 Homo sapiens cdna clone MAMMA1002805 5', mRNA
sequence.
ACCESSION AUI22649
VERSION AUI22649.1 GI:10947365
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 846)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
```

TITLE  
JOURNAL  
COMMENT

Isogai,T.  
HRI human cDNA project  
Unpublished (2000)  
Contact: Takao Isogai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: genomics@hri.co.jp  
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix  
Research Institute; cDNA library construction: Department of  
Virology, Institute of Medical Science, University of Tokyo, and  
Helix Research Institute.

FEATURES

source  
1..846  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MAMMA1002805"  
/tissue type="mammary gland"  
/clone\_lib="MAMMA1"  
/note="Vector: pME18SFL3"

ORIGIN

Query Match 35.2%; Score 768.4; DB 1; Length 846;  
Best Local Similarity 98.4%; Pred. No. 5.6e-192;  
Matches 797; Conservative 0; Mismatches 11; Indels 2; Gaps 2;  
QY 1144 AATGAATATAAAGGAATGCCAGGAGGAAGGAATCTTTATAAAATCCAACTTATTGACTAT 1203  
Db 1 AATGAATATAAAGGAATGCCAGGAGGAAGGAATCTTTATAAAATCCAACTTATTGACTAT 60  
QY 1204 ACAAAGTGACATCCCTCAGTTGTGAGTGAATCCGGAAGGAGTGTGAGTACTATTCTGTG 1263  
Db 61 ACAAAGTGACATCCCTCAGTTGTGAGTGAATCCGGAAGGAGTGTGAGTACTATTCTGTG 120  
QY 1264 TCATTTCAGTAAAGGCGGAAGTATTATCAGCTGAGATGTTCCGGTCTCTGGTCTGCCCCCTC 1323  
Db 121 TCATTTCAGTAAAGGCGGAAGTATTATCAGCTGAGATGTTCCGGTCTCTGGTCTGCCCCCTC 180  
QY 1324 TATATCTTACACAGCAGCGTGAATGATAAGGCGTGAGAGTCTCTGGAGAGACAAATTCAGCT 1383  
Db 181 TATATCTTACACAGCAGCGTGAATGATAAGGCGTGAGAGTCTCTGGAGAGACAAATTCAGCT 240  
QY 1384 TTGATATAAATCTGCGAGATGTTCCAGATGCCCTCCAAATAAATCGGACTTCATTTATTTTG 1443  
Db 241 TTGATATAAATCTGCGAGATGTTCCAGATGCCCTCCAAATAAATCGGACTTCATTTATTTTG 300  
QY 1444 AATGAACAATAATTTTGGTATCAGATGATCTTGCTCTCTCATTTTGTATAAATCCAAGAAA 1503  
Db 301 AATGAACAATAATTTTGGTATCAGATGATCTTGCTCTCTCATTTTGTATAAATCCAAGAAA 360  
QY 1504 TATCCTCTATATTTAGATGTTATGAGGCCCAATGTAGTCAAAAAGCAGACACTGTCTTC 1563  
Db 361 TATCCTCTATATTTAGATGTTATGAGGCCCAATGTAGTCAAAAAGCAGACACTGTCTTC 420  
QY 1564 AGACTGAATCGGGCCACTTACCTTGCAGACACAGAAAACATTTATAGTAGCTAGCTTTGAT 1623  
Db 421 AGACTGAATCGGGCCACTTACCTTGCAGACACAGAAAACATTTATAGTAGCTAGCTTTGAT 480  
QY 1624 GGCAGAGGAAGTGGTTACCAAGGAGATAAGATCATGATGCAATCAACAGAGACTGGGA 1683  
Db 481 GGCAGAGGAAGTGGTTACCAAGGAGATAAGATCATGATGCAATCAACAGAGACTGGGA 540  
QY 1684 ACATTTGAAGTTGAAGATCAAAATTTGAAGCGCCAGACAAATTTTCAAAAATGGGATTTGTG 1743  
Db 541 ACATTTGAAGTTGAAGATCAAAATTTGAAGCGCCAGACAAATTTTCAAAAATGGGATTTGTG 600  
QY 1744 GACAAACAAAGATTTGCAATTTTGGGCTGCTCATATGAGGGTACGTAACTCAATGGTC 1803  
Db 601 GACAAACAAAGATTTGCAATTTTGGGCTGCTCATATGAGGGTACGTAACTCAATGGTC 660

1804	CTGGATCGGAAAGTGGCGTGTTCAGTGTGGAAATAGCCGTGGCGCTGTATCCCAGTGG	1863
Qy		
661	CTGGATCGGAAAGTGGCGTGTTCAGTGTGGAAATACCCGTGGCGCTGTATCCCAGTGG	720
Db		
1864	GAGTACTATGACTCAGTGTATACAGAAAGCTTTACATGGGTCTCCCAACTCAGAAAGACAAC	1923
Qy		
721	GAAGTCTTACTCAATGT-CACAGAAAGCTTACATGGGTCTTCCAACCTCAGAAAGACAAC	779
Db		
1924	CTTGACCA-TTACAGAAATTTCAACAGTCAT	1952
Qy		
780	CTTGACCAATTTCCGGAATTTCAACAGTCAT	809
Db		

RESULT 13  
BQ937942 linear EST 21-AUG-2002  
LOCUS 900 bp mRNA  
ACCESSION 8922297 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:6470282  
DEFINITION 5'. mRNA sequence.

ACCESSION	BQ937942
VERSION	BQ937942.1
KEYWORDS	GI:22353420
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
	Hominidae; Homo.
REFERENCE	1 (bases 1 to 900)
AUTHORS	NIH-NCG <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

Email: [cqapbs-rc@mail.nih.gov](mailto:cqapbs-rc@mail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAMA4000 Row: m Column: 03  
High quality sequence stop: 639.

FEATURES	source
Location/Qualifiers	1. .900
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="IMAGE:6470282"
	/tissue_type="leiomyosarcoma"
	/lab_host="DH10B (phage-resistant)"
	/clone_lib="NIH_MGC_71"
	/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 2.1 kb."

Query Match 33.1%; Score 722; DB 5; Length 900;  
Best Local Similarity 99.4%; Pred. No. 1.1e-179;  
Matches 750; Conservative 0; Mismatches 10; Indels 2; Gaps 2;  
ORIGIN

Qy	1294	CTGAGATGTTCCGGTCTCGTGCTGCCCTCTATACTCTACACAGCAGCGGTGAATGATAAA	1353
Db	89	CTGTGCCTTTCCAGTCTCGTGCTGCCCTCTATACTCTACACAGCAGCGGTGAATGATAAA	148
Qy	1354	GGGCTGAGAGTCTCGGAAGACAATTCACTTTGGATAAATGCTGCAGAATGTCACCATG	1413
Db	149	GGGCTGAGAGTCTCGGAAGACAATTCACTTTGGATAAATGCTGCAGAATGTCACCATG	208
Qy	1414	CCCTCCAAAAAATCGGACTTCATTATTTTGAATGAAAACAAAATTTTCGTATCAGATGATC	1473
Db	209	CCCTCCAAAAAATCGGACTTCATTATTTTGAATGAAAACAAAATTTTCGTATCAGATGATC	268
Qy	1474	TTGCCTCCTCATTTTGTATAAATCAAGAAATATCCTCTACTATTAGATGTGTATGCAGGC	1533

269	Db	TTGCTCTCCATTTTTCATATAAATCCGAAGAATATCCCTCTACTATTAGATGTGTATGAGGC	328
1534	Qy	CCATGTAGTCAAAAAGCAGACACTGTCTTTCAGACTGAACCTGGGCCACTTACCTTGCAAAGC	1593
329	Db	CCATGTAGTCAAAAAGCAGACACTGTCTTTCAGACTGAACCTGGGCCACTTACCTTGCAAAGC	388
1594	Qy	ACAGAAAACATTATATAGTAGCTAGCTTTTGATGTGGCAGAGGAAGTGGTTTACCAAGAGAGATTAAG	1653
389	Db	ACAGAAAACATTATATAGTAGCTAGCTTTTGATGTGGCAGAGGAAGTGGTTTACCAAGAGAGATTAAG	448
1654	Qy	ATCATCGATCGCAATCAACAGAGACTGGGAACATTTGAAGTTTCAAGATCAAAATTGAAGCA	1713
449	Db	ATCATCGATCGCAATCAACAGAGACTGGGAACATTTGAAGTTTCAAGATCAAAATTGAAGCA	508
1714	Qy	GCAGACAATTTTCAAAAATGGGATTTGTGGCAACAACAAACGAAATTGCAATTTTGGGGCTGG	1773
509	Db	GCAGACAATTTTCAAAAATGGGATTTGTGGCAACAACAAACGAAATTGCAATTTTGGGGCTGG	568
1774	Qy	TCATATGGAGGGTACGTAACTCAATGGTCTCTGGGATCGGGAAGTGGCGTGTTCAGTGT	1833
569	Db	TCATATGGAGGGTACGTAACTCAATGGTCTCTGGGATCGGGAAGTGGCGTGTTCAGTGT	628
1834	Qy	GGAAATAGCGTGGCGCTGTATCCCGGTGGAGTACTATGACTCAGTGTACACAGAACGT	1893
629	Db	GGAAATAGCGTGGCGCTGTATCCCGGTGGAGTACTATGACTCAGTGTACACAGAACGT	688
1894	Qy	TACATGGGTCTCCCAACTCCAGAAGACAACCTTTGACCATTACAGAAATTCACACAGTCATG	1953
689	Db	TACATGGGTCTCCCAACTCCAGAAGACAACCTTTGACCATTACAGAAATTCACACAGTCATG	748
1954	Qy	AGCAGAGCTGAAATTTTAAAACAGTTGAGTAGCTCTTATTTTCATGGAAACAGCAGATGA	2012
749	Db	AGCAGAGCTGGAATTTTAAAACAGTTGAGTAGCTCTTATTTTCATGGAAACAGCAGATGA	808
2013	Qy	TAACGTTTAC - TTTACGAGCTCAGCTCAGATCTCCAAAGCCC	2053
809	Db	TAACGTTTCACTTTTCAGCAGTACGCTCAAAATCTCCCAAGGCC	850

RESULT	14
AL553051	
LOCUS	897 bp mRNA linear EST 30-MAR-2004
DEFINITION	AL553051 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CS0D1072YM19 5-PRIME, mRNA sequence.
ACCESSION	AL553051
VERSION	AL553051.3 GI:45857821
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	Li W.B., Gruber C., Jessee J. and Polayes D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	On Feb 15, 2001 this sequence version replaced gi:34274865.

**CONTACT:** Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremlieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primer  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoR V sites of the PCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 1082.r  
 For more information about this cluster, see  
<http://www.genoscope.cns.fr/cdna?s=CSDDI072AG10QP1&c=1082.r>.  
**FEATURES:**  
 Location/Qualifiers  
 1. .897  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"





Db 488 CAGTCATGAGCAGAGCTGAAATTTTAAACAAGTTGAGTACCTCCTTATTTCATGGAACAG 547  
Qy 2006 CAGATGATAACGTTCACTTTTCAGCAGTCAGCTCAGATCTCCAAAGCCCTGGTCGATGTTG 2065  
Db 548 CAGATGATAACGTTCACTTTTCAGCAGTCAGCTCAGATCTCCAAAGCCCTGGTCGATGTTG 607  
Qy 2066 GAGTGGATTTCCAGGCAATGTGTATACTGATGAAGACCATTGGAATAGCTAGCAGCACAG 2125  
Db 608 GAGTGGATTTCCAGGCAATGTGTATACTGATGAAGACCATTGGAATAGCTAGCAGCACAG 667  
Qy 2126 CACACCAACATATATATACCCACATGAGCCACTTCATAAAACAATGTTTCTCTTTACCT 2184  
Db 668 CACACCAACATATATATACCCACATGAGCCACTTCATAAAACAATGTTTCTCTTTACCT 726

Search completed: February 17, 2006, 07:21:17  
Job time : 8810 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2006, 05:06:42 ; Search time 654 Seconds  
(without alignments)  
7088.778 Million cell updates/sec

Title: US-10-659-055-2

Perfect score: 2184

Sequence: 1 agtcgaaactacactct.....aacatgtttctttacct 2184

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7204252 seqs, 1061369211 residues

Total number of hits satisfying chosen parameters: 14408504

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA\_New:  
1: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2182.4	99.9	2301	8	US-10-522-789-1
2	1654.6	75.8	4852	12	US-11-136-527-2130
3	751	34.4	2814	9	US-11-245-147-168
4	751	34.4	2814	12	US-11-186-284-54
5	119.8	5.5	524	6	US-09-925-065A-71343
6	111.2	5.1	2819	12	US-11-136-527-2716
7	101.4	4.6	615	6	US-09-925-065A-899711
8	101.4	4.6	615	6	US-09-925-065A-866910
9	77.8	3.6	2649	12	US-11-151-601-21
10	77.8	3.6	3143	12	US-11-151-601-19
11	70.4	3.2	379	12	US-11-128-061-396
12	70.4	3.2	379	12	US-11-128-061-4038
13	70.4	3.2	379	12	US-11-128-049-396
14	70.4	3.2	379	12	US-11-128-049-4038
15	62.8	2.9	1464	6	US-09-925-065A-546409
16	54.4	2.5	634	6	US-09-925-065A-735709
17	43.2	2.0	125594	8	US-10-658-986-5
18	42.6	2.0	597	6	US-09-925-065A-106709
19	42	1.9	1790	6	US-09-925-065A-956
20	41.6	1.9	575	6	US-09-925-065A-858506

c 21	41.6	1.9	578	6	US-09-925-065A-857835	Sequence 857835,
c 22	40.8	1.9	577	6	US-09-925-065A-474460	Sequence 474460,
c 23	40.8	1.9	1568	8	US-10-750-185-62484	Sequence 62484, A
c 24	40.8	1.9	1568	8	US-10-750-623-62484	Sequence 62484, A
c 25	40.4	1.8	600	12	US-11-136-527-6812	Sequence 6812, Ap
c 26	40.4	1.8	1400	12	US-11-136-527-6364	Sequence 6364, Ap
c 27	40.4	1.8	2567	12	US-11-136-527-2268	Sequence 2268, Ap
c 28	39.8	1.8	628	6	US-09-925-065A-752885	Sequence 752885,
c 29	39.8	1.8	867	8	US-10-750-185-62362	Sequence 62362, A
c 30	39.8	1.8	867	8	US-10-750-623-62362	Sequence 62362, A
c 31	39.4	1.8	629	8	US-10-750-185-37439	Sequence 37439, A
c 32	39.4	1.8	629	8	US-10-750-623-37439	Sequence 37439, A
c 33	39	1.8	618	6	US-09-925-065A-884672	Sequence 884672,
c 34	39	1.8	152335	12	US-11-121-086-73	Sequence 73, Appl
c 35	38.8	1.8	574	6	US-09-925-065A-263454	Sequence 263454,
c 36	38.8	1.8	643	6	US-09-925-065A-950355	Sequence 950355,
c 37	38.8	1.8	1271	8	US-10-750-185-62485	Sequence 62485, A
c 38	38.8	1.8	1271	8	US-10-750-623-62485	Sequence 62485, A
c 39	38.4	1.8	574	6	US-09-925-065A-263451	Sequence 263451,
c 40	38.4	1.8	574	6	US-09-925-065A-263452	Sequence 263452,
c 41	38.4	1.8	574	6	US-09-925-065A-263453	Sequence 263453,
c 42	38.4	1.8	580	6	US-09-925-065A-404802	Sequence 404802,
c 43	38.4	1.8	610	6	US-09-925-065A-210331	Sequence 210331,
c 44	38.4	1.8	633	6	US-09-925-065A-323095	Sequence 323095,
c 45	38	1.7	101786	12	US-11-117-187-199	Sequence 199, App

#### ALIGNMENTS

##### RESULT 1

US-10-522-789-1

; Sequence 1, Application US/10522789

; Publication No. US20050260732A1

; GENERAL INFORMATION:

; APPLICANT: TANABE SEIYAKU CO., LTD.

; TITLE OF INVENTION: Three-dimensional structure of dipeptidyl peptidase IV

; FILE REFERENCE: 03-039-PCT

; CURRENT APPLICATION NUMBER: US/10/522,789

; CURRENT FILING DATE: 2005-01-28

; PRIOR APPLICATION NUMBER: US 60/398,761

; PRIOR FILING DATE: 2002-07-29

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1

; LENGTH: 2301

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(2301)

; OTHER INFORMATION:

US-10-522-789-1

Query Match 99.9%; Score 2182.4; DB 8; Length 2301;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2183; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	AGTCGAAACTTACACTCTAACTGATTAATAAATACTTATAGACTGAAGTTATAC	60
Db	115	AGTCGAAACTTACACTCTAACTGATTAATAAATACTTATAGACTGAAGTTATAC	174
Qy	61	TCCTTAGATGGATTTTCAGATCATGATATCTCTACAAACAAGAAATAATATCTTGTA	120
Db	175	TCCTTAGATGGATTTTCAGATCATGATATCTCTACAAACAAGAAATAATATCTTGTA	234
Qy	121	TTCAATCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT	180
Db	235	TTCAATCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT	294
Qy	181	GGACATTTCTATCAATGATTAATCAATATCTCTGATGGCAGTTTATCTTTAGATAC	240
Db	295	GGACATTTCTATCAATGATTAATCAATATCTCTGATGGCAGTTTATCTTTAGATAC	354

QY 241 AACTAGTGAAGCAATGGAGGCAATCTTACACAGCTTCATATGACATTTTATGATTTAAAT 300  
 Db |||||  
 QY 355 AACTAGTGAAGCAATGGAGGCAATCTTACACAGCTTCATATGACATTTTATGATTTAAAT 414  
 Db |||||  
 QY 301 AAAAGGCAAGCTGATACAGAGAGAGGATTCAAAACAACACACAGTGGGTCAATGGTCA 360  
 Db |||||  
 QY 415 AAAAGGCAAGCTGATACAGAGAGAGGATTCAAAACAACACACAGTGGGTCAATGGTCA 474  
 Db |||||  
 QY 361 CAAGTGGGTCAATAATGGGCATATGTTGGAAACAATGACATTTATGTTAAATTTGAACCA 420  
 Db |||||  
 QY 475 CAAGTGGGTCAATAATGGGCATATGTTGGAAACAATGACATTTATGTTAAATTTGAACCA 534  
 Db |||||  
 QY 421 AATTTACCAAGTTTACAGAAATACATGAGCGGGAAGAAGATATAATAATGGAATA 480  
 Db |||||  
 QY 535 AATTTACCAAGTTTACAGAAATACATGAGCGGGAAGAAGATATAATAATGGAATA 594  
 Db |||||  
 QY 481 ACTGACTGGGTTTATGAAGAGGAAGTCTTCAGTGGCTACTCTGCTCTGFGGTGCTCCA 540  
 Db |||||  
 QY 595 ACTGACTGGGTTTATGAAGAGGAAGTCTTCAGTGGCTACTCTGCTCTGFGGTGCTCCA 654  
 Db |||||  
 QY 541 AACGGCACTTTTATGACATATGCCCAATTTAAGGACACAGAGTCCCAGCTTATTTGAATAC 600  
 Db |||||  
 QY 655 AACGGCACTTTTATGACATATGCCCAATTTAAGGACACAGAGTCCCAGCTTATTTGAATAC 714  
 Db |||||  
 QY 601 TCCCTTCTACTCTGATGAGTCACTGACAGTACCCAAAGACTGTACGGGTCCATATCCAAAG 660  
 Db |||||  
 QY 715 TCCCTTCTACTCTGATGAGTCACTGACAGTACCCAAAGACTGTACGGGTCCATATCCAAAG 774  
 Db |||||  
 QY 661 GCAGGAGCTGTGAATCCAACTGTAAAGTTCTTTGTTGTAATAACAGACTCTCTCAGCTCA 720  
 Db |||||  
 QY 775 GCAGGAGCTGTGAATCCAACTGTAAAGTTCTTTGTTGTAATAACAGACTCTCTCAGCTCA 834  
 Db |||||  
 QY 721 GTCAACAATGCCAATCTCCATACAAATCACTGCTCTGCTTCTATGTTGATAGGGATCAC 780  
 Db |||||  
 QY 835 GTCAACAATGCCAATCTCCATACAAATCACTGCTCTGCTTCTATGTTGATAGGGATCAC 894  
 Db |||||  
 QY 781 TACTTGTGTGATGATGATGGGCAACACAAAGAAAGAAATTTCTTGTGAGTGGCTCAGGAGG 840  
 Db |||||  
 QY 895 TACTTGTGTGATGATGATGGGCAACACAAAGAAAGAAATTTCTTGTGAGTGGCTCAGGAGG 954  
 Db |||||  
 QY 841 ATTCAAGAACTATTGGTCTATGATATTTGTGACTATGATGAATCCAGTGGGAAGATGGAAC 900  
 Db |||||  
 QY 955 ATTCAAGAACTATTGGTCTATGATATTTGTGACTATGATGAATCCAGTGGGAAGATGGAAC 1014  
 Db |||||  
 QY 901 TGTCTAGTGGCAAGGCAACACATTTGAATGATGATCTACTGGCTGGGTGGAAGATTTAGG 960  
 Db |||||  
 QY 1015 TGTCTAGTGGCAAGGCAACACATTTGAATGATGATCTACTGGCTGGGTGGAAGATTTAGG 1074  
 Db |||||  
 QY 961 CCTTCAGAACTCATTTTACCCTTGATGATGATGATGATCTTCAAGATCATCAGCAATGAA 1020  
 Db |||||  
 QY 1075 CCTTCAGAACTCATTTTACCCTTGATGATGATGATGATCTTCAAGATCATCAGCAATGAA 1134  
 Db |||||  
 QY 1021 GAAGGTTTACAGACACATTTGCTATTTTCCAAATAGATAAAAAAGACTGCACATTTATTA 1080  
 Db |||||  
 QY 1135 GAAGGTTTACAGACACATTTGCTATTTTCCAAATAGATAAAAAAGACTGCACATTTATTA 1194  
 Db |||||  
 QY 1081 AAGGCACTGGGAAGTCACTGGGATAGAACTCTAAACAGATGATTTATCTATCTACTACT 1140  
 Db |||||  
 QY 1195 AAGGCACTGGGGAAGTCACTGGGATAGAACTCTAAACAGATGATTTATCTATCTACTACT 1254  
 Db |||||  
 QY 1141 AGTAATGAATATAAAGGAATGCCAGGAGGAAGAAATCTTTTAAATAATCCAATTTATGAC 1200  
 Db |||||  
 QY 1255 AGTAATGAATATAAAGGAATGCCAGGAGGAAGAAATCTTTTAAATAATCCAATTTAGTAC 1314  
 Db |||||  
 QY 1201 TATACAAAAGTGACATGCTCCTCAGTGTGAGCTGAATCCGGAAGAGTGTGAGTACTATTCT 1260  
 Db |||||  
 QY 1315 TATACAAAAGTGACATGCTCCTCAGTGTGAGCTGAATCCGGAAGAGTGTGAGTACTATTCT 1374  
 Db |||||  
 QY 1261 GTGTCAATTCAGTAAAGGCGGAAGATTTATCAGCTGAGATGTTCCGGTCTGCTGTCGCC 1320  
 Db |||||  
 QY 1375 GTGTCAATTCAGTAAAGGCGGAAGATTTATCAGCTGAGATGTTCCGGTCTGCTGTCGCC 1434  
 Db |||||

QY 1321 CTCATATCTCTACACAGCAGCGTGAATGATATAAAGGGCTGAGAGTCTCTGGAAGACAATTC 1380  
 Db |||||  
 QY 1435 CTCATATCTCTACACAGCAGCGTGAATGATATAAAGGGCTGAGAGTCTCTGGAAGACAATTC 1494  
 Db |||||  
 QY 1381 GCTTTGGATATAAATGCTCCAGAAATGTCAGATGCCCTCCAAAATACTGGACTTCAATTT 1440  
 Db |||||  
 QY 1495 GCTTTGGATATAAATGCTCCAGAAATGTCAGATGCCCTCCAAAATACTGGACTTCAATTT 1554  
 Db |||||  
 QY 1441 TTGAATGAAAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGTATAAATCCAAG 1500  
 Db |||||  
 QY 1555 TTGAATGAAAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGTATAAATCCAAG 1614  
 Db |||||  
 QY 1501 AAATATCTCTCTACTATTAAGATGTTATGACAGGCCCATGTAGTCAAAAAGCAGACACTGTC 1560  
 Db |||||  
 QY 1615 AAATATCTCTCTACTATTAAGATGTTATGACAGGCCCATGTAGTCAAAAAGCAGACACTGTC 1674  
 Db |||||  
 QY 1561 TTCAGACTGAATCTGGGCCACTTACCTTGCAGACACAGAAACATTTATAGTACTAGCTTTT 1620  
 Db |||||  
 QY 1675 TTCAGACTGAATCTGGGCCACTTACCTTGCAGACACAGAAACATTTATAGTACTAGCTTTT 1734  
 Db |||||  
 QY 1621 GATGGCAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGCAATGCAATCAACAGAGACTG 1680  
 Db |||||  
 QY 1735 GATGGCAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGCAATGCAATCAACAGAGACTG 1794  
 Db |||||  
 QY 1681 GGAACTTTGAAGTTGAAGATCAAAATGAAGCAGCAGACAAATTTTCAAAAATGGGATTT 1740  
 Db |||||  
 QY 1795 GGAACTTTGAAGTTGAAGATCAAAATGAAGCAGCAGACAAATTTTCAAAAATGGGATTT 1854  
 Db |||||  
 QY 1741 GTGGCAACAACGAATTTGGGGCTGTTCAATGATGGAAGGTACGTAACTCAATG 1800  
 Db |||||  
 QY 1855 GTGGCAACAACAAGATTTGGGGCTGTTCAATGATGGAAGGTACGTAACTCAATG 1914  
 Db |||||  
 QY 1801 GTCTTGGGATCGGGAAGTGGGCTGTTCAAGTGTGGAATAGCGCTGCGCTCTGATCCCGG 1860  
 Db |||||  
 QY 1915 GTCTTGGGATCGGGAAGTGGGCTGTTCAAGTGTGGAATAGCGCTGCGCTCTGATCCCGG 1974  
 Db |||||  
 QY 1861 TGGGAGTACTATGACTCAGTGTACACAAAGTGTACATGGGTCTCCCAACTCCAGAAGAC 1920  
 Db |||||  
 QY 1975 TGGGAGTACTATGACTCAGTGTACACAAAGTGTACATGGGTCTCCCAACTCCAGAAGAC 2034  
 Db |||||  
 QY 1921 AACCTTGACCATTAACAGAAATTCACAGTCAAGCAGCAGAGCTGGAATTTTAAACAAGTT 1980  
 Db |||||  
 QY 2035 AACCTTGACCATTAACAGAAATTCACAGTCAAGCAGCAGAGCTGGAATTTTAAACAAGTT 2094  
 Db |||||  
 QY 1981 GAGTACCTCTCTTATTCATGGAACACAGATGATAACGTTTCACTTTTCAAGCAGTCAAG 2040  
 Db |||||  
 QY 2095 GAGTACCTCTCTTATTCATGGAACACAGATGATAACGTTTCACTTTTCAAGCAGTCAAG 2154  
 Db |||||  
 QY 2041 ATCTCCAAAGCCCTGGTCTGATGTTGGAGTGGATTTCCAGGCAATGTGGTATCTGATGAA 2100  
 Db |||||  
 QY 2155 ATCTCCAAAGCCCTGGTCTGATGTTGGAGTGGATTTCCAGGCAATGTGGTATCTGATGAA 2214  
 Db |||||  
 QY 2101 GACCATGGAATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2160  
 Db |||||  
 QY 2215 GACCATGGAATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2274  
 Db |||||  
 QY 2161 ATAAACCAATGTTCTCTTTTACCT 2184  
 Db |||||  
 QY 2275 ATAAACCAATGTTCTCTTTTACCT 2298  
 Db |||||

RESULT 2

US-11-136-527-2130  
 ; Sequence 2130, Application US/11136527  
 ; Publication No. US20050287570A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wyeth  
 ; APPLICANT: Mounts, William M  
 ; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes  
 ; FILE REFERENCE: 031896-041000 (AM101086)  
 ; CURRENT APPLICATION NUMBER: US/11/136,527  
 ; CURRENT FILING DATE: 2005-05-25  
 ; PRIOR APPLICATION NUMBER: US 60/574,294

; PRIOR FILING DATE: 2005-05-26  
 ; NUMBER OF SEQ ID NOS: 362830  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 2130  
 ; LENGTH: 4852  
 ; TYPE: DNA  
 ; ORGANISM: Rattus norvegicus  
 US-11-136-527-2130

Query Match 75.8%; Score 1654.6; DB 12; Length 4852;  
 Best Local Similarity 84.7%; Pred. No. 0;  
 Matches 1850; Conservative 16; Mismatches 310; Indels 9; Gaps 1;  
 7 AAAACTTACCTCTAATGATTTACTTAAATAAATACTTATAGACTGAAGTTACTCTTAA 66  
 203 AGAAGCTTACACTAKCTGACTATTTAAAGAAATACCTTTCCGGGTCAAGTCTACTCTCTG 262  
 67 AGATGGATTTGATCATGATATCTCTACAAACAGAGAAATATATCTTGTGATTTCAAT 126  
 263 CGGTGGGTTTCAGATTCTGATACCTCTACAGCAAGAAACAAATATCTTGTCTATTTCAAT 322  
 127 GCTGAATATGAAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTTGGACAT 186  
 323 GCTGACACAGGAAACAGCTCCATTTCTTGGAGAACAGTACCTTTGAGATCTTTGGAGAT 382  
 187 TCTATCAATGATTTATTCATATCTCTGATGGGCAAGTTTATTTCTCTTGAATATACAACTAC 246  
 383 TCTATAGTATTTATTCAGTGTCCACGACAGACTGTTTCTTCTTCTAGATATACAAATAT 442  
 247 GTGAAGCAATGGAGCAATCTCTACACAGCTTCATATGATCATTTATGATTTAAATAAAGG 306  
 443 GTGAAGCAATGGAGCAATCTCTACACAGCTTCATATGATTTATGATTTAAATAAAGG 502  
 307 CAGCTGATTTACAGAGAGAGATTTCCAAACACACACAGTGGTCCACATGTTCCACAGTG 366  
 503 CAGCTGATTTACAGAGAGAGATTTCCAAATATACACAGTGGATCACATGTTCCAAAGAA 562  
 367 GGTCAATAAATGGCATATGTTTGGAAACATGACATTTATGTTAAATTTGAACCAATTTA 426  
 563 GGTCAATAAATGGCATATGTTTGGAAACATGATTTATGTTAAATTTGAACCAATTTG 622  
 427 CCAAGTTACAGAAATACATGAGCGGGGAAAGAAATATATATATATATATATATATATAT 486  
 623 CCTAGTCAWAGGAYCACATCAACAGAAAGAAATGTATTTAAACGGWATAAATGAC 682  
 487 TGGGTTTATGAGAGAGATCTTCTAGTCCCTACTCTCTCTGTTGTTGTTCTCCAAACGGC 546  
 683 TGGGTTTATGAGAGAGAAATCTTCCGTTGCTTACTCCGCACTGTTGTTGTTCTCCAAACGGC 742  
 547 ACTTTTATGATATGCTCAATTTTAAACGACAGAGATCCCACTTATTTGATATCTCTCTTC 606  
 743 ACTTTTATGATATGCTCAATTTTAAACGACAGAGATCCCACTTATTTGATATCTCTCTTC 802  
 607 TACTCTGATGATGATCACTGATGATCCCAAGATCTGTACGGGTTCCATATCCAAAGGAGGA 666  
 803 TACTCTGATGATGATCACTGATGATCCCAAGATCTGTGATTTCCGTTACCCAAAGGAGGA 862  
 667 GCTGTGAATCACTGATGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 726  
 863 GCTGTGAATCACTGATGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 922  
 727 AATGCAATCTTCATACAAATCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 786  
 923 ACTAGATTTCCATGCAATCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 982  
 787 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 846  
 983 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1042  
 847 AACTATTGGTTCATGATTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 906  
 1043 AACTATTGGTTCATGATTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1102

QY 907 GTGGCA CGGCAACACATTTGAAATGAGTACTACTGGCTGGTGTGGAAGATTTAGGCTTCA 966  
 DB 1103 ACGACG CAGAGCATATTTGAAACGAGTGCACAGGCTGGKCGGAAGATTTAGGCTTCA 1162  
 QY 967 GAACTCATTTTACCTTTGATGTTAAATAGCTTCTTACAGATCATCAGCAATGAAAGAT 1026  
 DB 1163 GAACCCACATTTCCCTCCGACGGAAGCAGCTTCTATAAAATCGTCAAGATGAGC 1222  
 QY 1027 TACAGACATTTTCTATTTTCCAAATAGATA-----AAAAAGACTGCACTTTATTT 1077  
 DB 1223 TACAAACATCTGCCAGTTCAGAAAGATAGGAAACCCGAAACAGGCTGTATCATTTAT 1282  
 QY 1078 ACAAAGGACCTCGGAAGTCTATCGGATAGAGCTCTAAACAGTATTTATCTATCTAC 1137  
 DB 1283 ACAAAGGAGCTCGGAAGTCTATTTAGTATCGAAGCTCTGACAGCGATTTATCTGACTAC 1342  
 QY 1138 ATTAGTATGATATTAAGGAATGCCAGGAGAGAGATCTTTTATAAAATCCAACTTTAT 1197  
 DB 1343 ATTAGTATGATATTAAGGAATGCCAGGAGAGAGATCTTTTATAAAATCCAACTTTAT 1402  
 QY 1198 GACTATACAAAGTGCATGCTCTGATTTGTGAGCTGAAATCCGAAAGGTTGTCAGTACTAT 1257  
 DB 1403 GACCAACAAATAGAGAGTGCCTTGTGAGCTGAAATCCAGAAAGATGTCAGATTTAT 1462  
 QY 1258 TCTGTGCTATTTCAGTAAAGGCGGAGATTTATCAGCTGAGATGTTCCGGTCTCTGGTCTG 1317  
 DB 1463 TGGGTGCTATTCTTAAAGAGGCAAGTACTATCAGCTGGGATGCGCGGCGCTCTGGTCTG 1522  
 QY 1318 CCCCTCTATCTCTACACAGAGCGTGAATGATTAAGGGCTGAGAGTCTCTGGAAGACAT 1377  
 DB 1523 CCCCTCTACCTCTGCACTCGCAGCACTGATCAAAAGAGCTGAGAGTCTCTGGAGCAAT 1582  
 QY 1378 TCAGCTTTGGAATAAATGCTGCAAGATGCTCCAGATGCTCCCAAAAGCTGAGCTTTCA 1437  
 DB 1583 TCTGCTTTGATTAATAATGCTGCAAGATGCTCAAAATGCTTCAAAAGATTTGAGCTTTCA 1642  
 QY 1438 ATTTTGAATGAACAAATTTTGGTATCAGATGATCTTGGCTCTCTCAATTTTGATAAATCC 1497  
 DB 1643 GTTCTGAATGAACAAAGATTTTGGTATCAATGATCTTACCTCTCAATTTTGATAAATCC 1702  
 QY 1498 AAGAAATATCTCTATCTATTTAGATGTTGATGACGGCCCATGTAGTCAAAAGAGACAT 1557  
 DB 1703 AAGAAATATCTCTATCTATTTAGATGTTGATGACGGCTCTCTGAGTCAAAAGAGACAT 1762  
 QY 1558 GTCTTCAAGCTGACCTGCGGCACTTCTCTGCAAGCAGAGAAACATTTATGATGATGATG 1617  
 DB 1763 GCCTTCAAGTCAATCTGCGGCACTTCTCTGCAAGCAGAGAAACATCTATGATGATG 1822  
 QY 1618 TTTGATGGCAGAGAGATGTTTACCAAGGAGATGATGATGATGATGATGATGATGATGATG 1677  
 DB 1823 TTTGATGGCAGAGATGTTTACCAAGGAGATGATGATGATGATGATGATGATGATGATG 1882  
 QY 1678 CTGGGAACATTTGAAAGTGAAGATCAAAATGGAAGCAGCAGCAAAATTTTCAAAATGGA 1737  
 DB 1883 CTGGGAACATTTGAAAGTGAAGATCAAAATGGAAGCAGCAGCAGCAAAATTTTCAAAATGGA 1942  
 QY 1738 TTTTGGCAACAAACCAATTTGCAATTTTGGGCTGCTCATATGAGGGTGTAGTCACTCA 1797  
 DB 1943 TTTTGGCAACCAACCAATTTGCAATTTTGGGCTGCTCATATGAGGGTGTAGTCACTCA 2002  
 QY 1798 ATGCTCTGGATCGGGAATGCGGTGTTCAAGTGTGGAATGAGCGCTGGCGCTGTATGCC 1857  
 DB 2003 ATGCTCTGGATCGGGAATGCGGTGTTCAAGTGTGGAATGAGCGCTGGCGCTGTATGCC 2062  
 QY 1858 CGGTGGAGTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1917  
 DB 2063 CGGTGGAGTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2122  
 QY 1918 GACAACTTGCACATTTACAGAAATTTCAACAGTCTATGAGCAGAGCTGAAATTTTAAACAA 1977  
 DB 2123 GACAACTTGCACATTTACAGAAATTTCAACAGTCTATGAGCAGAGCTGAAATTTTAAACAA 2182  
 QY 1978 GTTGAGTACCTCTCTTATTTATGGAACAGAGATGATGATGATGATGATGATGATGATG 2037

Db 2193 GTTGAGTACCTCTTATTCACGGTACAGCAGATGATATGTTCACTTTTCAGCAGTCAGCT 2242  
 QY 2038 CAGATCTCCAAAGCCCTGGTTCGATGTTGGAGTGGATTTCCAGGCAATGTGGTATCTGAT 2097  
 Db 2243 CAGATCTCCAAAGCCCTGGTGGATGCTGGCGTGGATTTCCAGCAATGTGGTACACGGAC 2302  
 QY 2098 GAAGACCATGGAATAGTACGACGACGACGACCAACATATATATATACCCACATGAGCCAC 2157  
 Db 2303 GAAGACCATGGATCGCCAGCAGCAGCTACCCAGCAGCATCTATTCOCACATGAGCCAT 2362  
 QY 2158 TTCTATAAAACAAATGTTTCTCTTTAC 2182  
 Db 2363 TTCTCCAGCAGTGTCTCTCTTAC 2387

RESULT 3

US-11-245-147-168  
 ; Sequence 168, Application US/11245147  
 ; Publication No. US20060030541A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GARCIA, TERESA  
 ; APPLICANT: ROMAN ROMAN, SERGIO  
 ; APPLICANT: BARON, ROLAND  
 ; APPLICANT: CALL, KATHERINE  
 ; APPLICANT: THEILHABER, JOACHIM  
 ; APPLICANT: CONNOLLY, TIMOTHY  
 ; APPLICANT: JACKSON, AMANDA  
 ; APPLICANT: BUSHNELL, STEVEN  
 ; APPLICANT: RAWADI, GEORGES  
 ; TITLE OF INVENTION: GENES INVOLVED IN OSTEOGENESIS, AND METHODS OF USE  
 ; FILE REFERENCE: 37991-0023  
 ; CURRENT APPLICATION NUMBER: US/11/245,147  
 ; CURRENT FILING DATE: 2005-10-07  
 ; PRIOR APPLICATION NUMBER: PCT/IB02/02211  
 ; PRIOR FILING DATE: 2002-04-05  
 ; PRIOR APPLICATION NUMBER: 60/281,400  
 ; PRIOR FILING DATE: 2001-04-05  
 ; NUMBER OF SEQ ID NOS: 246  
 ; SOFTWARE: PatentIn Ver. 3.2  
 ; SEQ ID NO 168  
 ; LENGTH: 2814  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: Homo sapiens fibroblast activation protein, alpha  
 ; OTHER INFORMATION: (FAP), mRNA  
 US-11-245-147-168

Query Match 34.4%; Score 751; DB 9; Length 2814;  
 Best Local Similarity 60.9%; Pred. No. 1.3e-170;  
 Matches 1320; Conservative 0; Mismatches 825; Indels 24; Gaps 5;  
 QY 15 CACTCTAAGTACTTCTTAAATACTTATAGACTGAGTATCTCTTAAAGTGGAT 74  
 Db 334 CACACTGAAGATATTTTAAATGGAACATTTTCTTATAAAACATTTTTCCTCAACTGGAT 393  
 QY 75 TTCCAGATCATGAATATCTCTACAAACAGAAATATATCTTGGTATTCATCTGCTGAATA 134  
 Db 394 TTCAGGACAGATATCTTCATCATCTCTCAGTACATATAGTACTTTTATAATTTGA 453  
 QY 135 TGGAAACAGCTCAGTTTTCTTGGAGAACAGTACATTTTGATGATTTTGGACATCTATCAAA 194  
 Db 454 AACAGCAACATATATACCATTTTGGATTAAGAACCATGAAAGTGTGAATGCTTCAA- 512  
 QY 195 TGATTTTCAATATCTCTGATGGGAGTTTATCTCTTAGAATACAACTACGTGAAGCA 254  
 Db 513 --ATTACGGCTTATCACTGATCGGCAATTTGTATATCTAGAAAGTATTTTCAAGCT 570  
 QY 255 ATGGAGGCATTTCTACACAGCTTTCATATGACATTTTATGATTTTAAATAAAGGACGCTGAT 314  
 Db 571 TTGGAGATCTCTTACACAGCAACATATTTACATCTATGACCTTAGCAATGGAGAAATTTGT 630

QY 315 TACAGAGAGAGGATTTCCAAACAAACACACAGTGGGTACATGGTCCACAGTGGGTCAATA 374  
 Db 631 AAGAGGAAATCAGCTTCTCTGCTCCAAATTCAGTATTTATGCTGGTCGCTGTTGGGAGTAA 690  
 QY 375 ATTGGCATATGTTTGGAAACAATGACATTTATGTTAAATTTGAAACCAATTTTACCAAGTTA 434  
 Db 691 ATTAGCATATGCTATCAAAACAATATCTATTTGAAACAAAGACCCAGGAGATTCACCTTT 750  
 QY 435 CAGATCATCATGGAGGGGAAAGAGATATATATATATATATATATATATATATATATATAT 494  
 Db 751 TCAAAATTAACATTTAATGGAAGAGAAATATAATTTAATGGAATCCCGACAGCTGGGTATA 810  
 QY 495 TGAAGAGGAAGTCTTCAGTGCCTACTCTGCTCTGCTGGTGGTCTCCAAACGGCAGCTTTTT 554  
 Db 811 TGAAGAGGAATGCTTCTCAAAAATATGCTCTCTGGTGGTCTCTCTAATGGAATTTT 870  
 QY 555 AGCATATGCCCAATTTTAAACGACACAGAAAGTCCCACTTATTTGAATATCTCTCTTACTCTGA 614  
 Db 871 GGCATATGCGGAAATTTAATGATTAAGGATATACCAGTTATTCCTATTTCTATTATATGGCA 930  
 QY 615 TGAGTCACTGCGACTACCCAAAGACGTGTACGGGTCCATATCCAAAGGCGAGGCTGTGAA 674  
 Db 931 TGA-----ACAATATCTTAGAACAAATAATTTCCATACCCAAAGGCTGGAGCTTAAGAA 984  
 QY 675 TCCAACTGTAAAGTCTTTTGTGTAATAACAGACTCTCTCAGCTCAGTCACCAATGCCAAC 734  
 Db 985 TCCCGTTGTTGCGATATTTATTTATTCGATAC-----CACTTACCCTCGTATGTAGG 1035  
 QY 735 TTCCATACAAATCACTGCTCTCTGCTTCTATGTTGATAGGGGATCACTACTTGTGTGATGT 794  
 Db 1036 TCCCCAGGAAGTGCCTGTTCCAGCAATGATAGCCTCAAGTGAATATTATTTTCACTGGCT 1095  
 QY 795 GACATGGGCAACACAGAAAGAAATTTCTTTCAGTGGCTCAGGAGATTCAGAACTATTC 854  
 Db 1096 CACGTGGGTTTACTGATGAACGAGTATGTTTGCAGTGGCTTAAAGAGAGTCAGAAATGTTTC 1155  
 QY 855 GGTCAATGATATTTGTGACTATGATGAATCAAGTGAAGATGGAATCTGTTAGTGGCAG 914  
 Db 1156 GGTCTCTGTATATGTGACTTTCAGGGAAGACTGGCAGACATGGGATTTGTTCCAAAGACCA 1215  
 QY 915 GCACACATTTGAAATGAGTACTTACTGCTGGTTCGGAAGATTTAGGCTTTCAGAACCTCA 974  
 Db 1216 GGAGCATATAGAAGAAAGCAGAACTGGATGGCTGGTGGATTTCTTTTTCAGACCACT 1275  
 QY 975 TTTTACCCTTGATGTGTAATAGCTTCTCAAGATCATCAGCATGAAGGTTTACAGACA 1034  
 Db 1276 TTTTCAGCTATGATGCCATTTCTGACTACAAATATTTAGTGAACAGGATGGCTACAAACA 1335  
 QY 1035 CATTTGCTATTTCCAAATAGATATAAAGAGCTGCACATTTTATACAAAGGCACTCTGGA 1094  
 Db 1336 TATTCACATATATCAAGACACTGTGGAAATGCTATTCAATTTACAGTGGCAGTGGGA 1395  
 QY 1095 AGTCAATCGGGATAGAAGCTCTAACCAAGTGAATTTACTATCTACATTTAGTGAATGATATA 1154  
 Db 1396 GSCCATATATATTTTCAGAGTAAACACAGGATTCACATGTTTATTTCTAGCAATGAATTTCA 1455  
 QY 1155 AGGAATGCCAGGAGGAAGGATCTTTTAAATTCACATTTAGTACTA---TACAAAAGT 1211  
 Db 1456 AGAATACCTTGGAAAGAAACATCTACAGAAATGAGCTTGAAGCTATTCCTCCAGCAAA 1515  
 QY 1212 GACATGCTCAGTCTGAGCTGAATCCCGAAGGTGTCAGTACTATTCTGTGCTCATTCAG 1271  
 Db 1516 GAAGTGTGTACTTCCCATCTAAGGAAGAAAGGTGCCAATATTTACAGCAAGTTTCAG 1575  
 QY 1272 TAAAGAGGCGAGTATTTATCAGCTGAGATGTTTCGGTCTGTGCTGTGCCCCCTCTATCTCT 1331  
 Db 1576 CGACTACGCCAAGTACTATGCATTTGCTGTACGGCCCGGCACTCCCATTTTCCACCCT 1635  
 QY 1332 ACACAGCAGCGTAATGATTAAGGGCTGAGAGTCTCTGGNAGACAAATTCAGCTTTGGATA 1391  
 Db 1636 TCATGATGGACGCACTGATCAAGAAATTTAAATTCCTGGAAAGAAACAGGAATTTGAAA 1695  
 QY 1392 AATGCTGCAGAAATGTCAGATGCCCTCCAAAAAAGCTGGGACTTCAATTTATTTGAATGAAAC 1451



```
Db 1696 TGCTTGAATAATCCAGCTGCCCTAAAGAGAAATTAAGAACTTGAAGTAGATGAAT 1755
Qy 1452 AAAATTTTGGTATCAGATGATCTGCCTCCTCAATTTTGATTAATCAAGAAATATCCCTCT 1511
Db 1756 TACTTTATGATACAGATGATTTCTCCTCCTCAATTTGACAGATCAAGAGATATCCCTT 1815
Qy 1512 ACTATTAGATGTATGAGGCCCATGTAGTCAAAAGCAGACATGTCTTCAGACTGAA 1571
Db 1816 GCTAATTCAGGTATGTGGTGGTCCCTGCAGTCAGAGTGTAAAGGTCTGTATTGCTGTAA 1875
Qy 1572 CTGGSCCACTTACCTTTCAGCAGACAGAAACATATATAGTACGTCTTGTGGCAGAG 1631
Db 1876 TTGGATATCTTATCTTGCAAGTAAGAAAGGATGTCATTGCTTGGTGGATGGTCGAGG 1935
Qy 1632 AAGTGGTTTCAAGCAGATAGATCATGTCATGCAATCAACAGAGACTGGGAACATTTGA 1691
Db 1936 AACAGCTTTCAGGTGACAACTCTCTATGCACTGTATCGAAGCTGGGTGTTATGA 1995
Qy 1692 AGTTGAAGATCAAAATGGAAGCAGCAGACAAATTTTCAAAAATGGGATTTGGACACAA 1751
Db 1996 AGTTGAAGCAGATTAACAGCTGTTCAGAAATTCAGAAATGGGTTTCATTGATGAAA 2055
Qy 1752 ACAGATTGCAATTTGGGCTGTCTATATGAGGGGTACGTAACTCAATGGTCTGGGATC 1811
Db 2056 AAGATAGCCATATGGGCTGTCTATGAGGATACGTTTCATCACTGGCCCTTGCATC 2115
Qy 1812 GGGAGTGGCTGTTCAGATGTGGAATAGCCGTGGCGCTGTATCCCGTGGGATGACTA 1871
Db 2116 TGGAACTGCTCTTTTCAAAATGGTATAGCAGTGGCTCCAGTCTCCAGCTGGAAATTA 2175
Qy 1872 TGACTCAGTGTACACAGAACGTTACATGGGTCTCCCACTCCAGAAAGCAACCTTGACCA 1931
Db 2176 GCGCTCTGTACACAGAGATTCATGGGTCTCCCAACAAAGATGATATCTTGACCA 2235
Qy 1932 TTACAGAAATTCACAGTCATGACGACAGCTGAAATTTTAAACAGTTGAGTACCTCT 1991
Db 2236 CTATAAGAAATTCACAGTGGTGGCAAGAGCAGAAATATTTTCAGAAATGTAGACTATCTCT 2295
Qy 1992 TATTCATGGAACAGCAGATGATAAGCTTCACTTTTCAGCAGTCAGCTCAGATCTCCAAAGC 2051
Db 2296 CATCCAGGAACAGCAGATGATATGTCATCTTCAAACTCAGCAGATGTCTAAAGC 2355
Qy 2052 CTTGGTTCAGTGTGGAGTGGATTTCCAGGCAATGTGTATATCTGATGAAGACCATGGAAT 2111
Db 2356 TCTGGTTAATGACAAAGTGGATTTCCAGGCAATGTGTACTCTGACCAAGACCGGCTT 2415
Qy 2112 AGCTAGCAGCAGACACCAACATATATATATACCAATGAGCCACTTCATATAAACATG 2171
Db 2416 ATCCGGCTGTCCACGAAAC---CACTTATACACCAACATGACCCACTTCTTAAAGCAGTG 2472
Qy 2172 TTTCTCTTT 2180
Db 2473 TTTCTCTTT 2481
```

RESULT 4  
US-11-186-284-54  
; Sequence 54, Application US/11186284  
; Publication No. US20050266493A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Berger, Allison  
; APPLICANT: Guillemette, Tracy L.  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Monahan, John E.  
; APPLICANT: Thibodeau, Stephen N.  
; APPLICANT: Burgart, Lawrence J.  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF COLON CANCER  
; FILE REFERENCE: MPM01-029P2RNM

```
; CURRENT APPLICATION NUMBER: US/11/186,284  
; CURRENT FILING DATE: 2005-07-21  
; PRIOR APPLICATION NUMBER: US/10/301,822  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 60/339,971  
; PRIOR FILING DATE: 2001-12-10  
; PRIOR APPLICATION NUMBER: US 60/361,978  
; PRIOR FILING DATE: 2002-03-05  
; PRIOR APPLICATION NUMBER: US 60/381,988  
; PRIOR FILING DATE: 2002-05-20  
; NUMBER OF SEQ ID NOS: 228  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 54  
; LENGTH: 2814  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (209)...(2491)  
US-11-186-284-54  
  
Query Match 34.4%; Score 751; DB 12; Length 2814;  
Best Local Similarity 60.9%; Pred. No. 1.3e-170;  
Matches 1320; Conservative 0; Mismatches 825; Indels 24; Gaps 5;  
  
Qy 15 CACTCTAACTGATTAACCTTAAATAAATCTTATAGACTGAAGTTATATCTCTTAAAGATGGAT 74  
Db 334 CACACTGAAGATATTTTAAATGGAACATTTCTTATATAAACATTTTTCACAACTGGAT 393  
Qy 75 TTCAGATCATGAATATCTCTACAAAGAAATAATATCTTGGTATTCATATGCTGAATA 134  
Db 394 TTCAGGACAAAGATATCTTCAATCAATCTGCAGATAACAATATAGTACTTTATAATATGA 453  
Qy 135 TGGAAACAGCTCAGTTTCTTGGAGAACAGTACATTTTGTAGTAAAGTGTGAATGCTTCAA 194  
Db 454 AACAGGACAATCATATACCATTTTGGTAAATAGAACCAATGAAAGTGTGAATGCTTCAA 512  
Qy 195 TGATTTATTCATATCTCTGATGGCGAGTTTATTTCTTTAGAAATACAACTACGTGAAGCA 254  
Db 513 --ATTACGGCTTATCACCTGATCGCAATTTGTATATCTAGAAAGTATTTCAAGCT 570  
Qy 255 ATGAGGATTCCTACACAGCTTCATATGACATTTATGATTAATAAAGAGCACTGAT 314  
Db 571 TTGAGATATCTCTTACACAGCAACATATATCATCTATGACCTTAGCAATGGAGAAATTTGT 630  
Qy 315 TACAGAGAGAGGATTCACAAACACACACAGTGGGTACATGGTCCACAGTGGGTCTATAA 374  
Db 631 AAGAGGAAATGAGCTTCTCTGTCCTCAATTTTCAATTTATGCTGGTGGCTGTGGGAGTAA 690  
Qy 375 ATTGGCATATGTTTGGAAACATATGATCATTTATGTTAAATTTGAACCAAAATTTTACCAAGTTA 434  
Db 691 ATTAGCATATGCTTATCAAAACATATCTATTGAAACAAAGACAGAGATCCACCTTT 750  
Qy 435 CAGAAATCACAATGGACGGGAAAGAGATATATATATAATGAAATACTGACTGGGTTTA 494  
Db 751 TCAAAATAACATTTAATGGAAGAGAAATAAAATATTTAAATGGAATCCAGAGTGGGTTTA 810  
Qy 495 TGAAGAGGAGTCTTTCAGTGGCTACTCTGCTCTGGTGGTGGTCTCCAAAGCGCACTTTT 554  
Db 811 TGAAGAGGAAATGCTTCTCTCAAAATATGCTCTCTGGTGGTGGTCTCTTAAATGGAATAATTTT 870  
Qy 555 AGCATATGCCCAATTTTAAACGACAGAGAGTCCCACTTATTGAAATACCTCTTCTACTCTGA 614  
Db 871 GGCATATGGCGAATTTAATGATAGGATATACCAAGTATTCCTTATTTCTTATTTGGCGA 930  
Qy 615 TGAGTCACTGCAGTACCCAAAGACTGTACGGGTTTCCATATCCAAAGGAGCAGAGCTGTGAA 674  
Db 931 TGA-----ACAAATATCTAGAACAAATAAATAATTTTCCATATCCAAAGGCTGGAGCTAAGAA 984  
Qy 675 TCCAACTGTAAAGTTCTTGTGTTAAATACAGACTCTCTCAGCTCAGTCCAGTACCAATGCAAC 734  
Db 985 TCCCGTTGTTGGGATATTTTATTCGATAC-----CACTTACCTCGGTATGTAGG 1035
```

```
QY 735 TTCCATACAAATCACCTGCTCCTGCTTCTATGTTGATAGGGGATCACTACTGTGTGATGT 794
Db 1036 TCCCAAGGAAGTGCTGTTCCAGCAATGATAGCCTCAAGTGAATATTTATTTCAAGTTGGCT 1095
QY 795 GACATGGGCAACACAAAGAAGATTTCTTTGCAAGTGGCTCAGGAGGATTCAGAACTATTTC 854
Db 1096 CACGTGGGTACTGATGAACGAGTATGTTTGCAAGTGGCTTAAAGAGAGTCCAGAAATGTTTC 1155
QY 855 GGTCAATGGATATTTGATGATATGATGAATCCAGTGGAGATGGAATGCTGTAGTGGCAGG 914
Db 1156 GGTCTCTGTCTATATGTCATCTCAGGGAAGACTGGCAGACATGGGATTTGTCCAAGACCCA 1215
QY 915 GCACACATTTGAAATGAGTACTACTGGCTGGGTGGAGATTTAGSCCTTCAGAACTCA 974
Db 1216 GGAGCATATAGAAAGAACAGAACTGGATGGGCTGGTGGATTTCTTTGTTTCAAGACCAGT 1275
QY 975 TTTTACCCCTTGATGTTAATAGCTTCTACAAGATCATCAGCAATAGAAAGGTTTACAGACA 1034
Db 1276 TTTGAGTATGATGCCATTTTCGTACTACAAATATTTAGTGAAGGATGGCTACAAACA 1335
QY 1035 CATTTGCTATTTCCAAATAGATAAAGAACTGCACATTTTATACAAAGGACACCTGGGA 1094
Db 1336 TATTCACATATATCAAGACACATGTGGAAATGCTATTTCAAAATTACAAGTGGCAAGTGGGA 1395
QY 1095 AGTCATCGGGATAGAAAGCTCTAACCAAGTATTTATCTATACATTCATTAATAGTAATATAA 1154
Db 1396 GGCCATAAATATTTACAGAGTAAACACAGGATTCACGTGTTTATTTAGCAATGAATTTGA 1455
QY 1155 AGGAATGCGAGGAGGAAGAAATCTTTTATAAAATCCAACTTATTTGACTA---TACAAAAGT 1211
Db 1456 AGAATACCTCGGAAGAAGAAACATCTACAGAAATAGCAATGGAAGCTATCCTCCAAGCAA 1515
QY 1212 GACATGCCCTCAGTTGTGACGTGAATCCGGAAGGTGTAGTACTATTCTGTGTCATTGAC 1271
Db 1516 GAAGTGTGTTACTTGGCCATCTAAGGAAAGAAAGGTGCCAATATACACAGCAAGTTTCAG 1575
QY 1272 TAAAGAGGCGAAGTATTTATCAGCTGAGATGTTCCGGTCTGGTCTGCCCTCTCTACTCT 1331
Db 1576 CGACTACGCCAAGTACTATGCACTTGTCTGCTACGGCCACAGGATCCCCATTTCCACCT 1635
QY 1332 ACACAGCAGCGTGAATGATAAAGGGCTGAGAGTCTCTGGAAGACAATTCAGCTTTGGATTA 1391
Db 1636 TCATGATGAGCGCACTGATCAAGAAATTAATCTCTGGAAGAAACAAAGAAATTTGAAAA 1695
QY 1392 AATGCTGCAAGATGTCAGATGCCCTCCTCAAAAACCTGGACTCATTTATTTTGAATGAAC 1451
Db 1696 TGCTTTGAAAAATATCCAGCTGCTTAAAGAGGAAATTAAGAAACTTTGAAAGTAGATGAAT 1755
QY 1452 AAAATTTTGTGATCAGATGATCTGGCTCTCTCATTTTGTATAATCCAAGAAATATCCTCT 1511
Db 1756 TACTTTATGGTACAAGTGAATCTTCTCTCAATTTGACAGATCAAGAAAGTATCCCTT 1815
QY 1512 ACTATTAGATGTGTATGACAGGCCCATGTAGTCAAAAAGCAGACACTGTCTTCAGACTGAA 1571
Db 1816 GCTAATTCAGTGTATGTGTGCTCCCTGCAGTCAGAGTGAAGTCTGTATTTGCTGTTAA 1875
QY 1572 CTGGGCCACTTACTCTTGAAGCAACAGAAACATTATATAGTAGCTAGCTTTGATGGCAGAGG 1631
Db 1876 TTGGATATCTTATCTTTGCAAGTAAAGGAAGGATGGTCAATTGGCTTGGTGGTGGTGGAG 1935
QY 1632 AAGTGGTTACCAAGGAGATAAGATCATGATGCAATCAACAGAAAGCTGCGAATTTTCA 1691
Db 1936 AACAGCTTTCCAAAGGTGACAAACTCTCTATGCAAGTGTATCGAAAGCTGGGTGTTATGA 1995
QY 1692 AGTTGAAGATCAAAATTAAGACAGCCACAGACAATTTTCAAAAATGGGATTTTGTGACAA 1751
Db 1996 AGTTGAAGACCGAGATTACAGCTGTACAGAAATTCATAGAAATGGGTTTCATTGATGAAA 2055
QY 1752 ACGAATTTGCAATTTGGGGCTGGTCATATGAGGGTACGTAACTCAATGGTCTCGGATC 1811
Db 2056 AAGAATAGCCATATGGGGCTGGTCTTATGGAGGATACGTTTTCATCATCTGGCCCTTGCA 2115
QY 1812 GCGAAGTGGCGTGTTCAGAGTGTGGAATAGCCGTGGCGCTGTATCCCGGTGGGAGTACTA 1871
```

```
Db 2116 TGGAACTGGTCTTTTCAAAATGTGGTATAGCAGTGGCTCCAGTCTCCAGCTGGGAATATTA 2175
QY 1872 TGACTCAGTGTACACAGAAAGTGTATCATGGGTCTCCAACTCCAGAGACAACTTTGACCA 1931
Db 2176 CGCGTCTGTCTACACAGAGAGATTTCATGGGTCTCCCAACAAAGGATGATTAATCTTGAGCA 2235
QY 1932 TTACAGAAATTCACACAGTCAAGCAGAGCTGAAAAATTTTAAACAAGTTGAGTACTCTCT 1991
Db 2236 CTATAGAAATTCACCTGTGATGGCAAGACAGAAATATTTCAGAAATGTAGACTATCTTCT 2295
QY 1992 TATTTCATGGAACAGCAGATGATAACGTTTCACTTTTTCAGCAGTCAAGTCTCAGATCTCCAAAGC 2051
Db 2296 CATCACGGAACAGCAGATGATAATGTGCACCTTTCAAAACTCAGCAGATTTGCTTAAAGC 2355
QY 2052 CCGTGTCAATGTGGAGTGGATTTCCAGGCAATGTGGTATATCTGATGAAGACATGGAAT 2111
Db 2356 TCTGGTTAATGCACAAGTGGATTTCCAGGCAATGTGGTACTCTGACCAAGAACCAAGCTT 2415
QY 2112 AGCTAGCAGCAGCAGCAGCAGCAATATATATACCCACATAGGCCACTTCTATAAAACAATG 2171
Db 2416 ATCCGGCTGTCCAGAAC---CACTTATACCCACATAGCCACTTCTCTAAGCAGTG 2472
QY 2172 TTTCTCTTT 2180
Db 2473 TTTCTCTTT 2481
```

RESULT 5

```
US-09-925-065A-71343/c
; Sequence 71343, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 71343
; LENGTH: 524
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-71343
```

```
Query Match 5.5%; Score 119.8; DB 6; Length 524;
Best Local Similarity 88.4%; Pred. No. 6.4e-19;
Matches 130; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 645 GGTTCATATCCAAAGGAGGAGCTGTGAATCCAACTGTAAAGTTCTTTTGTGTAATAC 704
Db 498 GATTTTCTTCTAGGAGGAGCTGTGAATCCAACTGTAAAGTTCTTTTGTGTAATAC 439
QY 705 AGACTCTCTCAGCTCAGTCACCAATGCAACTTCCATACAAATCACTGCTCCTGCTTCTAT 764
Db 438 AGACTCTCTCAGCTCAGTCACCAATGCAACTTCCATACAAATCACTGCTCCTGCTTCTAT 379
QY 765 GTTGATAGGGATCACTACTTGTGTGA 791
Db 378 GTTGATAGGGTAAGACTCTTGTCTGGA 352
```



; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 866910  
; LENGTH: 615  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-866910

Query Match 4.6%; Score 101; DB 6; Length 615;  
Best Local Similarity 98.1%; Pred. No. 2.3e-14;  
Matches 101; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1771 TGGTCATATGAGGGTACGTAACTCAATGTCCTGGGATCGGGAAGTGGCGTGTCAAG 1830  
Db 478 TAGTCATATGAGGGTACGTAACTCAATGTCCTGGGATCGGGAAGTGGCGTGTCAAG 419  
QY 1831 TGTGAATAGCGTGGCGCTGTATCCCGTGGGAGTACTATG 1873  
Db 418 TGTGAATAGCGTGGCGCTGTATCCCGTGGGAGTATATG 376

## RESULT 9

US-11-151-601-21  
; Sequence 21, Application US/11151601  
; Publication No. US20060003413A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Meyers, Rachel E.  
; APPLICANT: Olandt, Peter J.  
; APPLICANT: Kapeller-Libermann, Rosana  
; APPLICANT: Curtis, Rory A. J.  
; APPLICANT: Williamson, Mark  
; APPLICANT: Weich, Nadine  
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE.  
; TITLE OF INVENTION: AND PROTEASE FAMILY MEMBERS AND USES THEREOF  
; FILE REFERENCE: MPI00-054PRCPIONDIVIM  
; CURRENT APPLICATION NUMBER: US/11/151,601  
; CURRENT FILING DATE: 2005-06-13  
; PRIOR APPLICATION NUMBER: US 10/170,789  
; PRIOR FILING DATE: 2002-06-13  
; PRIOR APPLICATION NUMBER: US 09/797,039  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: PCT/US01/06525  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: US 60/186,061  
; PRIOR FILING DATE: 2000-02-29  
; PRIOR APPLICATION NUMBER: US 09/882,166  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: PCT/US01/19269  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/212,078  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: US 09/934,406  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: PCT/US01/26052  
; PRIOR FILING DATE: 2001-08-21  
; PRIOR APPLICATION NUMBER: US 60/226,740

; PRIOR FILING DATE: 2000-08-21  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 21  
; LENGTH: 2649  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-151-601-21

Query Match 3.6%; Score 77.8; DB 12; Length 2649;  
Best Local Similarity 48.7%; Pred. No. 1.7e-08;  
Matches 242; Conservative 0; Mismatches 252; Indels 3; Gaps 1;  
QY 1619 TTGATGGCAGAGGAGTGGTTACCAAGGAGATAAGATCATGTCATGCAATCAACAGAGAC 1678  
Db 2057 TAGACACAGGGGATCCTGTCCACGAGGCTTAATTTGAAGGCCCTTTAATATATAAA 2116  
QY 1679 TGGGAACATTTGAAGTTGAAGATCAAAATTGAAGCAGCCAGA---CAATTTTCAAAATGG 1735  
Db 2117 TGGGTCAATAGAAATTGACGATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATAG 2176  
QY 1736 GATTTGTGGACAAACAAAGAAATTGCAATTTGGGGCTGGTCATATGAGGGTACGTAACT 1795  
Db 2177 ATTTCAATTGACTTAGATCGTGTGGGCATCCAGCGTGGTCTTATGGAGGATACCTCTCCC 2236  
QY 1796 CAATGTCCTGGATCGGGAAGTGGCGTGTCAAGTGTGGAATAGCCGTGGCGCTGTAT 1855  
Db 2237 TGATGGCAATTAATGACAGAGGTCAGATATCTTCAGGGTGTCTATTTGGGGCCCGCATCA 2296  
QY 1856 CCGGTGGGAGTACTATGACTCAGTGTACACAGAAACGTTTACATGGGTCTTCCAACTCCAG 1915  
Db 2297 CTCGTGGATCTTCTATGATACAGATACACGGAAGTTATATGGTCACCCCTGACCAGA 2356  
QY 1916 AAGCAACCTTGACCATACAGAAATTCACAGTCAAGAGAGAGCTGAAAATTTTAAAC 1975  
Db 2357 ATGAACAGGGCTATTACTTAGGATCTGTGGCATCAAGACAGAAAAGTTCCCTCTGAAC 2416  
QY 1976 AAGTTGAGTACTCTCTTATTCATGAAACAGCAGATAGTAACGTTTCACTTTTCAGCAGTCA 2035  
Db 2417 CAAATCGTTTACTGCTCTTACATGTTTCCCTGGATGAGAAATGTCATTTTGCACATACCA 2476  
QY 2036 CTCAGATCTCCAAAGCCCTGGTCCGATGTTGGAGTGGATTTCCAGGCAATGTGTATCTG 2095  
Db 2477 GTATATTACTGAGTTTCTTATGAGGGCTGGAAGCCATATGATTTACAGATCTATCCTC 2536  
QY 2096 ATGAAGACCATGGAATA 2112  
Db 2537 AGGAGAGACACAGCATA 2553

## RESULT 10

US-11-151-601-19  
; Sequence 19, Application US/11151601  
; Publication No. US20060003413A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Meyers, Rachel E.  
; APPLICANT: Olandt, Peter J.  
; APPLICANT: Kapeller-Libermann, Rosana  
; APPLICANT: Curtis, Rory A. J.  
; APPLICANT: Williamson, Mark  
; APPLICANT: Weich, Nadine  
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE.  
; TITLE OF INVENTION: AND PROTEASE FAMILY MEMBERS AND USES THEREOF  
; FILE REFERENCE: MPI00-054PRCPIONDIVIM  
; CURRENT APPLICATION NUMBER: US/11/151,601  
; CURRENT FILING DATE: 2005-06-13  
; PRIOR APPLICATION NUMBER: US 10/170,789  
; PRIOR FILING DATE: 2002-06-13  
; PRIOR APPLICATION NUMBER: US 09/797,039  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: PCT/US01/06525

```
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 3143
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (229)...(2874)
US-11-151-601-19

Query Match      3.6%; Score 77.8; DB 12; Length 3143;
Best Local Similarity 48.7%; Pred. No. 1.8e-08;
Matches 242; Conservative 0; Mismatches 25; Indels 3; Gaps 1;

Qy 1619 TTGATGGCAGAGGAAGTGGTTACCAAGGAGATAAGATCATGATGCAATCAACAGAGAC 1678
Db 2285 TAGACAAACAGGGATCTCTGTCCACGAGGGCTTAAATTTGAAGGCGCTTTAAATATAAAA 2344

Qy 1679 TGGGAACATTTGAAGTTGAAGATCAATTTGAAGCAGCAGA---CAATTTCAAAATGG 1735
Db 2345 TGGGTCAATAGAATTTGACATCAGGTGGAAGGACTCCAATATCTAGCTTCTCGATATG 2404

Qy 1736 GATTTGTGGACAACAAGATTGCAATTTGGGGCTGTGCATATGAGGGTACGTTACCT 1795
Db 2405 ATTTCAATGATTTAGATCTGTGGCATCCACGGCTGGTCTCTATGGAGGATACCTCTCCC 2464

Qy 1796 CAATGGTCTGGGATCGGGAAGTGGCGTGTTCAGAGTGTGGAATAGCCGTGGCGCTGTAT 1855
Db 2465 TGATGGCATTAAATCAGAGGTCAGATATCTTCAGGGTTGCTATTGCTGGGCCCCAGTCA 2524

Qy 1856 CCGGTGGAGTACTATGACTCAGTGTACACAGAACGTTACATGGGTCTCCCAACTCCAG 1915
Db 2525 CTCTGTGGATCTTCTATGATACAGGATACAGGAACGTTATATGGGTCAACCTGACCCAGA 2584

Qy 1916 AAGCAACCTTGACCATTTACAGAAATTCACAGTCATGAGCAGAGCTGAAATTTTAAAC 1975
Db 2585 ATGAACAGGCTATTACTTAGGATCTGTGGCATGCAAGCAAGAAAGTTCCCTCTGAAC 2644

Qy 1976 AAGTTGAGTACTCTCTTATTTCATGGAACAGCAGATGATAACGTTTCACCTTCAGCAGTCAG 2035
Db 2645 CAATCGTTTACTCTCTTACATGGTTCTCTGGATGAGATGTCATTTTGCATACCA 2704

Qy 2036 CTCAGATCTTCAAGCCCTGGTTCGATGTGGAGTGGATTTCCAGGCAATGTGGTATATCTG 2095
Db 2705 GTATATTACTGAGTTTCTTGTAGTGGGCTGGAAGGCCATATGATTTACAGATCTATCCTC 2764

Qy 2096 ATGAAGACCATGGAATA 2112
Db 2765 AGGAGACACAGCATATA 2781

RESULT 11
US-11-128-061-396
; Sequence 396, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4038
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Cricetulus griseus
US-11-128-061-4038

Query Match      3.2%; Score 70.4; DB 12; Length 379;
Best Local Similarity 92.5%; Pred. No. 4.4e-07;
Matches 74; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1945 ACAGTCATGACGACAGCTGAAATTTTAAACAGTTGAGTACCTCTTATTTCATGGAACA 2004
Db 2 ACAGTCATGACGACAGCTGAAATTTTAAACAGTTGAGTACCTCTTATTTCATGGAACA 61

Qy 2005 GCAGATGATAACGTTCACTT 2024
Db 62 GCAGATGAGTTCAGTT 81

RESULT 12
US-11-128-061-4038
; Sequence 4038, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Hann, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; CURRENT FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; PRIOR FILING DATE: 2004-05-11
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4038
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Cricetulus griseus
US-11-128-061-4038

Query Match      3.2%; Score 70.4; DB 12; Length 379;
Best Local Similarity 92.5%; Pred. No. 4.4e-07;
Matches 74; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1945 ACAGTCATGACGACAGCTGAAATTTTAAACAGTTGAGTACCTCTTATTTCATGGAACA 2004
Db 2 ACAGTCATGACGACAGCTGAAATTTTAAACAGTTGAGTACCTCTTATTTCATGGAACA 61

Qy 2005 GCAGATGATAACGTTCACTT 2024
```

Db 62 GCAGATGCTGAGTTTCAGTT 81  
 ||||| | | | | | | | |

RESULT 13  
 US-11-128-049-396  
 ; Sequence 396, Application US/11128049  
 ; Publication No. US20060010513A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Melville, Mark W.  
 ; APPLICANT: Charlebois, Timothy S.  
 ; APPLICANT: Mounts, William M.  
 ; APPLICANT: Hann, Louane E.  
 ; APPLICANT: Sinacore, Martin S.  
 ; APPLICANT: Leonard, Mark W.  
 ; APPLICANT: Brown, Eugene L.  
 ; APPLICANT: Miller, Christopher P.  
 ; TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR  
 ; FILE REFERENCE: 01997.027700  
 ; CURRENT APPLICATION NUMBER: US/11/128,049  
 ; CURRENT FILING DATE: 2005-05-11  
 ; PRIOR APPLICATION NUMBER: US 60/570,425  
 ; PRIOR FILING DATE: 2004-05-11  
 ; NUMBER OF SEQ ID NOS: 7285  
 ; SOFTWARE: PatentIn version 3.3  
 ; SEQ ID NO 396  
 ; LENGTH: 379  
 ; TYPE: DNA  
 ; ORGANISM: Cricetus griseus  
 US-11-128-049-396

Query Match 3.2%; Score 70.4; DB 12; Length 379;  
 Best Local Similarity 92.5%; Pred. No. 4.4e-07;  
 Matches 74; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 1945 ACAGTCATGACGAGCTGAAATTTTAAACAAGTTGAGTACCTCTCTTATTCATGGAACA 2004  
 Db 2 ACAGTCATGACGAGCTGAAATTTTAAACAAGTTGAGTACCTCTCTTATTCATGGAACA 61  
 ||||| | | | | | | | |

QY 2005 GCAGATGATAAGTTTCAGTT 2024  
 ||||| | | | | | | | |  
 Db 62 GCAGATGCTGAGTTTCAGTT 81

RESULT 14  
 US-11-128-049-4038  
 ; Sequence 4038, Application US/11128049  
 ; Publication No. US20060010513A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Melville, Mark W.  
 ; APPLICANT: Charlebois, Timothy S.  
 ; APPLICANT: Mounts, William M.  
 ; APPLICANT: Hann, Louane E.  
 ; APPLICANT: Sinacore, Martin S.  
 ; APPLICANT: Leonard, Mark W.  
 ; APPLICANT: Brown, Eugene L.  
 ; APPLICANT: Miller, Christopher P.  
 ; TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR  
 ; FILE REFERENCE: 01997.027700  
 ; CURRENT APPLICATION NUMBER: US/11/128,049  
 ; CURRENT FILING DATE: 2005-05-11  
 ; PRIOR APPLICATION NUMBER: US 60/570,425  
 ; PRIOR FILING DATE: 2004-05-11  
 ; NUMBER OF SEQ ID NOS: 7285  
 ; SOFTWARE: PatentIn version 3.3  
 ; SEQ ID NO 4038  
 ; LENGTH: 379  
 ; TYPE: DNA  
 ; ORGANISM: Cricetus griseus  
 US-11-128-049-4038

Query Match 3.2%; Score 70.4; DB 12; Length 379;  
 Best Local Similarity 92.5%; Pred. No. 4.4e-07;  
 Matches 74; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 1945 ACAGTCATGACGAGCTGAAATTTTAAACAAGTTGAGTACCTCTCTTATTCATGGAACA 2004  
 Db 2 ACAGTCATGACGAGCTGAAATTTTAAACAAGTTGAGTACCTCTCTTATTCATGGAACA 61  
 ||||| | | | | | | | |  
 QY 2005 GCAGATGATAAGTTTCAGTT 2024  
 ||||| | | | | | | | |  
 Db 62 GCAGATGCTGAGTTTCAGTT 81

RESULT 15  
 US-09-925-065A-546409/c  
 ; Sequence 546409, Application US/09925065A  
 ; Publication No. US20040181048A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, David G.  
 ; TITLE OF INVENTION: Identification and Mapping of Single  
 ; FILE REFERENCE: 108827.135  
 ; CURRENT APPLICATION NUMBER: US/09/925,065A  
 ; CURRENT FILING DATE: 2001-08-08  
 ; PRIOR APPLICATION NUMBER: US 60/243,096  
 ; PRIOR FILING DATE: 2000-10-24  
 ; PRIOR APPLICATION NUMBER: US 60/252,147  
 ; PRIOR FILING DATE: 2000-11-20  
 ; PRIOR APPLICATION NUMBER: US 60/250,092  
 ; PRIOR FILING DATE: 2000-11-30  
 ; PRIOR APPLICATION NUMBER: US 60/261,766  
 ; PRIOR FILING DATE: 2001-01-16  
 ; PRIOR APPLICATION NUMBER: US 60/289,846  
 ; PRIOR FILING DATE: 2001-05-09  
 ; NUMBER OF SEQ ID NOS: 957086  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 546409  
 ; LENGTH: 1464  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-925-065A-546409

Query Match 2.9%; Score 62.8; DB 6; Length 1464;  
 Best Local Similarity 69.7%; Pred. No. 5.3e-05;  
 Matches 85; Conservative 0; Mismatches 37; Indels 0; Gaps 0;  
 QY 1752 AGCAATTCGAATTTGGGGCTGCTATATGAGGGTACGTAACCTCAATGGTCTGGGATC 1811  
 Db 756 ACTCATAACAATGTCTCTACAGTCTCTATGGAGGATACGTTTCATCACTGGCCCTTGCATC 697  
 ||||| | | | | | | | |  
 QY 1812 GGGAGTGGCGTGTTCGAATAGCCGCTGCGCTGTATCCCGTGGGAGTACTA 1871  
 Db 696 TGGAACTGGTCTTTTCAATGTGATAGCAGTGGCTCCAGTCTCCAGCTGGGATATTA 637  
 ||||| | | | | | | | |  
 QY 1872 TG 1873  
 Db 636 CG 635

Search completed: February 17, 2006, 06:44:30  
 Job time : 656 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2006, 05:01:13 ; Search time 1790 Seconds  
(without alignments)  
10089.565 Million cell updates/sec

Title: US-10-659-055-2

Perfect score: 2184

Sequence: 1 agtcgcaaaactacactct.....aacatgtttcttacct 2184

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA\_Main:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2184	100.0	3407	5	US-10-002-593-5
2	2184	100.0	3407	5	US-10-165-603-6
3	2184	100.0	3407	6	US-10-423-714-5
4	2184	100.0	3407	6	US-10-295-027-921
5	2184	100.0	3407	7	US-10-794-899-40
6	2184	100.0	3407	10	US-11-041-674-5
7	2182.4	99.9	2301	6	US-10-295-027-589
8	2182.4	99.9	3445	9	US-10-952-459-17
9	2182.4	99.9	3445	9	US-10-956-157-741
10	2182.4	99.9	3445	9	US-10-631-467-42
11	2182.4	99.9	3913	9	US-10-887-553A-216
12	2180.8	99.9	2924	7	US-10-641-643-1026
13	1677.6	76.8	3316	9	US-10-631-467-969
14	1656.2	75.8	4835	3	US-08-917-800A-1570
15	1656.2	75.8	4835	5	US-10-165-603-5
16	1656.2	75.8	4835	7	US-10-794-899-39
17	754.2	34.5	2366	6	US-10-101-510-683
18	752.6	34.5	2788	6	US-10-269-909-22
19	752.6	34.5	2788	8	US-10-723-860-4170
20	752.6	34.5	2788	9	US-10-887-553A-235
21	752.6	34.5	2788	9	US-10-956-157-1251
22	752.6	34.5	3138	5	US-10-198-846-13171
23	751	34.4	2814	3	US-09-962-832-108

## ALIGNMENTS

## RESULT 1

; Sequence 5, Application US/10002593  
; Publication No. US20020137120A1  
; GENERAL INFORMATION:  
; APPLICANT: Vanderbilt University  
; APPLICANT: Brown, Nancy J.  
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTING ENZYME  
; FILE REFERENCE: INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA  
; CURRENT APPLICATION NUMBER: US/10/002,593  
; CURRENT FILING DATE: 2001-10-31  
; PRIOR APPLICATION NUMBER: 60/244,524  
; PRIOR FILING DATE: 2000-10-31  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 5  
; LENGTH: 3407  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-002-593-5

Query Match 100.0%; Score 2184; DB 5; Length 3407;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGTCGCAAACTTACACTCTAACTGATTAACAAAATCTTATAGACTGAAGTTATAC 60  
Db 190 AGTCGCAAACTTACACTCTAACTGATTAACAAAATCTTATAGACTGAAGTTATAC 249

Qy 61 TCCTTAGATGGATTTTCAGATCATGATATCTCTACAAACAAGAAATAATATCTTGGA 120  
Db 250 TCCTTAGATGGATTTTCAGATCATGATATCTCTACAAACAAGAAATAATATCTTGGA 309

Qy 121 TTCAATCTCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 180  
Db 310 TTCAATCTCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 369

Qy 181 GGACATTTCTATCAATGATTTATCAATATCTCTGATGGCAGTTTATCTCTTAGAATAC 240  
Db 370 GGACATTTCTATCAATGATTTATCAATATCTCTGATGGCAGTTTATCTCTTAGAATAC 429

Qy 241 AACTACGTGAAGCATGAGGCATTTCTACACAGCTTCATATGACATTTATGATTAAT 300  
Db 430 AACTACGTGAAGCATGAGGCATTTCTACACAGCTTCATATGACATTTATGATTAAT 489



QY 301 AAAAGCAGCTGATTACAGAGAGAGGATTCCAAACAAACACACAGTGGGTCAATGGTCA 360  
DB |||||  
QY 490 AAAAGCAGCTGATTACAGAGAGAGGATTCCAAACAAACACAGTGGGTCAATGGTCA 549  
DB |||||  
QY 361 CCAAGTGGGTGATATAATGTCATATGTTGGAAACAATGACATTTATGTTAAAAATTGAACCA 420  
DB |||||  
QY 550 CCAAGTGGGTGATATAATGTCATATGTTGGAAACAATGACATTTATGTTAAAAATTGAACCA 609  
DB |||||  
QY 421 AATTATCAAGTTACAGAAATCAATGACCGGGGAAAGAAAGATATATATATATATGGAATA 480  
DB |||||  
QY 610 AATTATCAAGTTACAGAAATCAATGACCGGGGAAAGAAAGATATATATATATGGAATA 669  
DB |||||  
QY 481 ACTGAGTGGGTTATGAAGAGAGAGTCTTCAAGTGCCTACTCTGCTCTGTTGGTCTTCA 540  
DB |||||  
QY 670 ACTGAGTGGGTTATGAAGAGAGAGTCTTCAAGTGCCTACTCTGCTCTGTTGGTCTTCA 729  
DB |||||  
QY 541 AACGGCAGCTTTTATGATATATGCCCAGTATTAACGACACAGAAAGTCCCACTTATGGAATAC 600  
DB |||||  
QY 730 AACGGCAGCTTTTATGATATATGCCCAGTATTAACGACACAGAAAGTCCCACTTATGGAATAC 789  
DB |||||  
QY 601 TCCTTCTACTGATGAGTCACTGACGATACCCAAAGACTGTACGGGTTCCATATCCAAAG 660  
DB |||||  
QY 790 TCCTTCTACTGATGAGTCACTGACGATACCCAAAGACTGTACGGGTTCCATATCCAAAG 849  
DB |||||  
QY 661 GCAGGAGCTGGAATCCAACTGTAAAGTTCTTGTGTAATAACAGACTCTCTCAGTCA 720  
DB |||||  
QY 850 GCAGGAGCTGGAATCCAACTGTAAAGTTCTTGTGTAATAACAGACTCTCTCAGTCA 909  
DB |||||  
QY 721 GTCCCAATGCAACTTCCATACAAATCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780  
DB |||||  
QY 910 GTCCCAATGCAACTTCCATACAAATCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 969  
DB |||||  
QY 781 TACTGTGTGATGATGAGTCACTGAGGCAACAAGAAAGAAATTTCTTGTGAGTGGGTCA 840  
DB |||||  
QY 970 TACTGTGTGATGATGAGTCACTGAGGCAACAAGAAAGAAATTTCTTGTGAGTGGGTCA 1029  
DB |||||  
QY 841 ATTCCAGAACTATTCCGTGATGATATTTGATGATGATGATGATGATGATGATGATGATG 900  
DB |||||  
QY 1030 ATTCCAGAACTATTCCGTGATGATATTTGATGATGATGATGATGATGATGATGATGATG 1089  
DB |||||  
QY 901 TGCCTAGTGGGCAACACATTTGAAATGAGTACTACTGCTGCTGCTGCTGCTGCTGCTGCTG 960  
DB |||||  
QY 1090 TGCCTAGTGGGCAACACATTTGAAATGAGTACTACTGCTGCTGCTGCTGCTGCTGCTGCTG 1149  
DB |||||  
QY 961 CCTTCAGAACTCAATTTTACCTTGTATGATGATGATGATGATGATGATGATGATGATGATG 1020  
DB |||||  
QY 1150 CCTTCAGAACTCAATTTTACCTTGTATGATGATGATGATGATGATGATGATGATGATGATG 1209  
DB |||||  
QY 1021 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATAAAAAGACTGACATTTATTACA 1080  
DB |||||  
QY 1210 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATAAAAAGACTGACATTTATTACA 1269  
DB |||||  
QY 1081 AAAGGCACCTGGGAGTCACTGGGATGAGAGTCTTAACAGTGAATATCTATCTACTACATTT 1140  
DB |||||  
QY 1270 AAAGGCACCTGGGAGTCACTGGGATGAGAGTCTTAACAGTGAATATCTATCTACTACATTT 1329  
DB |||||  
QY 1141 AGTAATGAATATAAAGGAATGCGAGGAGGAATCTTTATAAATCCAACTTATTGAC 1200  
DB |||||  
QY 1330 AGTAATGAATATAAAGGAATGCGAGGAGGAATCTTTATAAATCCAACTTATTGAC 1389  
DB |||||  
QY 1201 TATACAAAGTGCATGCTCAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGT 1260  
DB |||||  
QY 1390 TATACAAAGTGCATGCTCAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGT 1449  
DB |||||  
QY 1261 GTGTCAATTCAGTAAAGAGCGAAGTATTTATGAGTGAATGTTCCGGTCTGCTGCTGCTGCTG 1320  
DB |||||  
QY 1450 GTGTCAATTCAGTAAAGAGCGAAGTATTTATGAGTGAATGTTCCGGTCTGCTGCTGCTGCTG 1509  
DB |||||  
QY 1321 CTCCTATCTACACAGCAGCGTGAATGATTAAGGCTGAGAGTCTCTGGAAGCAATTTCA 1380  
DB |||||  
QY 1510 CTCCTATCTACACAGCAGCGTGAATGATTAAGGCTGAGAGTCTCTGGAAGCAATTTCA 1569  
DB |||||  
QY 1381 GCTTTGGATAAATGCTGCAGAAATGTCAGATGCCCTCCAAAAAATCGGACTTCAATTATT 1440

DB 1570 GCTTTGGATAAAAATGCTGCAGAAATGCTGCAGATGCCCTCAAAAAAATCGACTTCAATTATT 1629  
QY 1441 TTGAATGAAACAAAATTTTGGTATCAGATGATCTTGGCTCCTCATTTTGGATTAATCCAG 1500  
DB 1630 TTGAATGAAACAAAATTTTGGTATCAGATGATCTTGGCTCCTCATTTTGGATTAATCCAG 1689  
QY 1501 AAATATCTCTACTACTATTAGATGATGATGAGGCCCATGTTAGTCAAAAAAGCAGACACTGTC 1560  
DB 1690 AAATATCTCTACTACTATTAGATGATGATGAGGCCCATGTTAGTCAAAAAAGCAGACACTGTC 1749  
QY 1561 TTGAGACTGAACTGGGCCCACTTTACCTTGCAGACACAGAAAAATTTATAGTACTAGCTTTT 1620  
DB 1750 TTGAGACTGAACTGGGCCCACTTTACCTTGCAGACACAGAAAAATTTATAGTACTAGCTTTT 1809  
QY 1621 GATGCGAGAGAGTGGTTTACCAAGAGATTAAGATCATGCTGCAATCAACAGAGACTG 1680  
DB 1810 GATGCGAGAGAGTGGTTTACCAAGAGATTAAGATCATGCTGCAATCAACAGAGACTG 1869  
QY 1681 GGAACATTTGAGTTGAAGATCAAAATTTGAAGAGTCAAAATTTGAAGAGTCAAAATTTGAAGAGT 1740  
DB 1870 GGAACATTTGAGTTGAAGATCAAAATTTGAAGAGTCAAAATTTGAAGAGTCAAAATTTGAAGAGT 1929  
QY 1741 GTGGACAAACAAACGAATTTGCGGCTGGTCAATTTGAGGGGTACGTAAACCTCAATG 1800  
DB 1930 GTGGACAAACAAACGAATTTGCGGCTGGTCAATTTGAGGGGTACGTAAACCTCAATG 1989  
QY 1801 GTCTGGGATCGGGAAGTGGCTGTTCAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG 1860  
DB 1990 GTCTGGGATCGGGAAGTGGCTGTTCAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG 2049  
QY 1861 TGGGAGTACTATGACTCAGTGPACACAGAACTTACATGGGTCTCCCAACTCCAGAAAGAC 1920  
DB 2050 TGGGAGTACTATGACTCAGTGPACACAGAACTTACATGGGTCTCCCAACTCCAGAAAGAC 2109  
QY 1921 AACCTTGACCAATTACAGAAATTTCAACAGTCAAGCAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1980  
DB 2110 AACCTTGACCAATTACAGAAATTTCAACAGTCAAGCAGTCAAGTCAAGTCAAGTCAAGTCAAGT 2169  
QY 1981 GAGTACCTCTTATTCATGGAACAGCAGATGATAGTTCACCTTTCAGCAGTCAAGTCAAGTCAAGT 2040  
DB 2170 GAGTACCTCTTATTCATGGAACAGCAGATGATAGTTCACCTTTCAGCAGTCAAGTCAAGTCAAGT 2229  
QY 2041 ATCTCCAAAGCCCTGCTGATGTTGGAGTGGATTTCCAGCAATGTTGATATCTGATGAA 2100  
DB 2230 ATCTCCAAAGCCCTGCTGATGTTGGAGTGGATTTCCAGCAATGTTGATATCTGATGAA 2289  
QY 2101 GACCATGGAATAGTACGACGACAGCAACCAACATATATATACCCATGAGCCACTTC 2160  
DB 2290 GACCATGGAATAGTACGACGACAGCAACCAACATATATATACCCATGAGCCACTTC 2349  
QY 2161 ATAAAAAATGTTCTCTTTACCT 2184  
DB 2350 ATAAAAAATGTTCTCTTTACCT 2373

## RESULT 2

US-10-165-603-6  
; Sequence 6, Application US/10165603  
; Publication No. US20030021792A1  
; GENERAL INFORMATION:  
; APPLICANT: Roben, Paul W.  
; TITLE OF INVENTION: TISSUE-SPECIFIC ENDOTHELIAL MEMBRANE  
; TITLE OF INVENTION: TISSUE-SPECIFIC ENDOTHELIAL MEMBRANE  
; FILE REFERENCE: TPTech.001A  
; CURRENT APPLICATION NUMBER: US/10/165,603  
; CURRENT FILING DATE: 2002-06-07  
; PRIOR APPLICATION NUMBER: 60/297,021  
; PRIOR FILING DATE: 2001-06-08  
; PRIOR APPLICATION NUMBER: 60/305,117  
; PRIOR FILING DATE: 2001-07-12  
; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6  
; LENGTH: 3407  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (76)...(2376)  
US-10-165-603-6

Query Match 100.0%; Score 2184; DB 5; Length 3407;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	AGTGGCAAACTTACACCTCTAACTGATTACTTAAATAATCTTATAGACTGAAGTTATAC	60
Db	190	AGTGGCAAACTTACACCTCTAACTGATTACTTAAATAATCTTATAGACTGAAGTTATAC	249
Qy	61	TCCTTAAGATGGATTTCAGATCATGAATATCTCTACAAACAGAAATAATATCTTGGTA	120
Db	250	TCCTTAAGATGGATTTCAGATCATGAATATCTCTACAAACAGAAATAATATCTTGGTA	309
Qy	121	TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT	180
Db	310	TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT	369
Qy	181	GGACATTTCTAATGATTATTTCAATATCTCTGATGGGCAAGTTTATTTCTTTAGAAATAC	240
Db	370	GGACATTTCTAATGATTATTTCAATATCTCTGATGGGCAAGTTTATTTCTTTAGAAATAC	429
Qy	241	AACCTAGTGAAGCAATGGAGGCATTTCTTACACAGCTTCATATGACATTTATGATTTAAAT	300
Db	430	AACCTAGTGAAGCAATGGAGGCATTTCTTACACAGCTTCATATGACATTTATGATTTAAAT	489
Qy	301	AAAAGGCAAGCTGATTACAGAGAGAGGATTTCCAAACAAACACACAGTGGGTCAATGTCAC	360
Db	490	AAAAGGCAAGCTGATTACAGAGAGAGGATTTCCAAACAAACACACAGTGGGTCAATGTCAC	549
Qy	361	CCAGTGGGTCAATAATTTGGCATATGTTTGGAAACAATGACATTTATGTTTAAATTTGAACCA	420
Db	550	CCAGTGGGTCAATAATTTGGCATATGTTTGGAAACAATGACATTTATGTTTAAATTTGAACCA	609
Qy	421	AAATTTACCAAGTTACAGAAATCACATGGAACGGGAAAGAGATATAATATATAATTTGAATA	480
Db	610	AAATTTACCAAGTTACAGAAATCACATGGAACGGGAAAGAGATATAATATATAATTTGAATA	669
Qy	481	ACTGACTGGGTTTATGAAGAGAGAGTCTTCACTGCTACTCTGCTCTGCTGCTGCTCTCCA	540
Db	670	ACTGACTGGGTTTATGAAGAGAGAGTCTTCACTGCTACTCTGCTCTGCTGCTGCTCTCCA	729
Qy	541	AACGGCACTTTTATGACATATGCCCAATTTAAACGACAGAGTCCCACTTTATTTGAATAC	600
Db	730	AACGGCACTTTTATGACATATGCCCAATTTAAACGACAGAGTCCCACTTTATTTGAATAC	789
Qy	601	TCCTTTCTACTCTGATGATGATCTGACAGTACCCAAAGAGCTGACGGGTTCCATATCCAAAG	660
Db	790	TCCTTTCTACTCTGATGATGATCTGACAGTACCCAAAGAGCTGACGGGTTCCATATCCAAAG	849
Qy	661	GCAGAGCTGGAATCCAACTGTGAAGTTCTTTGTTGTTAAATAACAGACTCTCTCAGCTCA	720
Db	850	GCAGAGCTGGAATCCAACTGTGAAGTTCTTTGTTGTTAAATAACAGACTCTCTCAGCTCA	909
Qy	721	GTCAACCAATGCAATTCATACAAATCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	780
Db	910	GTCAACCAATGCAATTCATACAAATCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	969
Qy	781	TACTTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	840
Db	970	TACTTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1029
Qy	841	ATTGAGAACTATTCGGTTCATGGATATTTGATGATGATGATGATGATGATGATGATGATGAT	900
Db	1030	ATTGAGAACTATTCGGTTCATGGATATTTGATGATGATGATGATGATGATGATGATGATGAT	1089

Qy	901	TGCTTAGTGGCAACGCAACATTTGAATGAGTACTACTGCTGGTGGTGGAGATTTAGG	960
Db	1090	TGCTTAGTGGCAACGCAACATTTGAATGAGTACTACTGCTGGTGGTGGAGATTTAGG	1149
Qy	961	CTTTGAGAACTCTCAATTTTACCCCTTGGATGGTAAATAGCTTCTTCAAGATCATCAGCAATGAA	1020
Db	1150	CTTTGAGAACTCTCAATTTTACCCCTTGGATGGTAAATAGCTTCTTCAAGATCATCAGCAATGAA	1209
Qy	1021	GAAGGTTACAGACACATTTGCTATTTCCAAATAGATATAAAGAGTGCACATTTATATACA	1080
Db	1210	GAAGGTTACAGACACATTTGCTATTTCCAAATAGATATAAAGAGTGCACATTTATATACA	1269
Qy	1081	AAAGGCACTCTGGGAAGTCACTCGGATAGAAAGCTCTAAACAGTGAATATCTATCTATCTACAT	1140
Db	1270	AAAGGCACTCTGGGAAGTCACTCGGATAGAAAGCTCTAAACAGTGAATATCTATCTATCTACAT	1329
Qy	1141	AGTAATGAATATAAAGGAATGCCAGGAGGAAGGAATCTTTTATAAAATCCAACTTATTTGAC	1200
Db	1330	AGTAATGAATATAAAGGAATGCCAGGAGGAAGGAATCTTTTATAAAATCCAACTTATTTGAC	1389
Qy	1201	TATACAAAGCTGACATGCTCAGTTGTGAGCTGAATCGGAAAGGTGTCTAGTACTATTTCT	1260
Db	1390	TATACAAAGCTGACATGCTCAGTTGTGAGCTGAATCGGAAAGGTGTCTAGTACTATTTCT	1449
Qy	1261	GTGTCAATTAAGAGAGGCGAAGTATTTATCAGCTGAGATGTTCCGGTCTCTGGTCTGCCC	1320
Db	1450	GTGTCAATTAAGAGAGGCGAAGTATTTATCAGCTGAGATGTTCCGGTCTCTGGTCTGCCC	1509
Qy	1321	CTCTATATCTTACACAGCAGCGTGAATGATAAAGGCTGAGAGTCTCTGGAAAGACAAATTTCA	1380
Db	1510	CTCTATATCTTACACAGCAGCGTGAATGATAAAGGCTGAGAGTCTCTGGAAAGACAAATTTCA	1569
Qy	1381	GCTTTGATTAATGCTGTCAGAAATGTCAGATGCTCCAGTCCCTCCAAATACTGGACTTCAATTT	1440
Db	1570	GCTTTGATTAATGCTGTCAGAAATGTCAGATGCTCCAGTCCCTCCAAATACTGGACTTCAATTT	1629
Qy	1441	TTGAATCAAAACAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGTATTAATTTCCAA	1500
Db	1630	TTGAATCAAAACAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGTATTAATTTCCAA	1689
Qy	1501	AAATATCTCTATATTTAGATGTTATGCAAGGCTGAGAGTCTCTGAAAGACAGACTGTC	1560
Db	1690	AAATATCTCTATATTTAGATGTTATGCAAGGCTGAGAGTCTCTGAAAGACAGACTGTC	1749
Qy	1561	TTCAAGCTGAACTGGGCACTTACCTTTGCAAGACACAGAAACATTTATAGTACTAGCTTT	1620
Db	1750	TTCAAGCTGAACTGGGCACTTACCTTTGCAAGACACAGAAACATTTATAGTACTAGCTTT	1809
Qy	1621	GATGGCAGAGAGAGTGGTTTACCAAGGAGATAAGATCATGATGCAATCAACAGAAAGACTG	1680
Db	1810	GATGGCAGAGAGAGTGGTTTACCAAGGAGATAAGATCATGATGCAATCAACAGAAAGACTG	1869
Qy	1681	GGAACTTTGAAAGTTGAAGATCAAAATTTGAAGCAGCAGACAAATTTTCAAAATGGGATTT	1740
Db	1870	GGAACTTTGAAAGTTGAAGATCAAAATTTGAAGCAGCAGACAAATTTTCAAAATGGGATTT	1929
Qy	1741	GTGGCAACAAACGAATTTGGGGCTGGTTCATATGGAGGGTACGTAACTCAATG	1800
Db	1930	GTGGCAACAAACGAATTTGGGGCTGGTTCATATGGAGGGTACGTAACTCAATG	1989
Qy	1801	GTCTCTGGATCGGGAGTGGCTGTTCAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG	1860
Db	1990	GTCTCTGGATCGGGAGTGGCTGTTCAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG	2049
Qy	1861	TGGGAGTACTATGACTCAGTGTACACAGAAACGTTTACATGGGTCTCCCAACTCCAGAAAGAC	1920
Db	2050	TGGGAGTACTATGACTCAGTGTACACAGAAACGTTTACATGGGTCTCCCAACTCCAGAAAGAC	2109
Qy	1921	AACCTTGACCAATTTACAGAAATTTCAAGTCAATGAGCAGAGCTGAAATTTTAAACAGTT	1980
Db	2110	AACCTTGACCAATTTACAGAAATTTCAAGTCAATGAGCAGAGCTGAAATTTTAAACAGTT	2169









Db 490 AAAAGGCGAGTGAATTAAGAGAGAGAGGATTCAAACAAACACACAGTGGGTCAATGGTCA 549  
 Qy 361 CCAGTGGGTCAATAATTTGGCATATGTTTGGCAACATGACATTTATGTTAAATTTGAACCA 420  
 Db 550 CCAGTGGGTCAATAATTTGGCATATGTTTGGCAACATGACATTTATGTTAAATTTGAACCA 609  
 Qy 421 AATTTACCAAGTTACAGAAATCACATGGACGGGGAAAGAAAGATATAAATATAAATGGAATA 480  
 Db 610 AATTTACCAAGTTACAGAAATCACATGGACGGGGAAAGAAAGATATAAATATAAATGGAATA 669  
 Qy 481 ACTGACTGGGTTTATGAAGAGAGAGTCTTTCAGTGGCTACTCTGCTGTGGTGGTCTCCA 540  
 Db 670 ACTGACTGGGTTTATGAAGAGAGAGTCTTTCAGTGGCTACTCTGCTGTGGTGGTCTCCA 729  
 Qy 541 AACGGCACTTTTATGACATATGCCCAATTTAAACAGACACAGAAAGTCCCACTTATTGAATAC 600  
 Db 730 AACGGCACTTTTATGACATATGCCCAATTTAAACAGACACAGAAAGTCCCACTTATTGAATAC 789  
 Qy 601 TCCTTTCTACTCTGATGAGTCACTGCAGTACCACAAAGACTGTACCGGGTTCCATATCCAAAG 660  
 Db 790 TCCTTTCTACTCTGATGAGTCACTGCAGTACCACAAAGACTGTACCGGGTTCCATATCCAAAG 849  
 Qy 661 GCAGAGCTGTGAATCCCACTGTAAAGTTCCTTGTGTAAATACAGACTCTCTCAGTCA 720  
 Db 850 GCAGAGCTGTGAATCCCACTGTAAAGTTCCTTGTGTAAATACAGACTCTCTCAGTCA 909  
 Qy 721 GTCAACCAATGCACTTCCATACAAATCACTGCTCTGCTCTTATGTTGATAGGGGATCAC 780  
 Db 910 GTCAACCAATGCACTTCCATACAAATCACTGCTCTGCTCTTATGTTGATAGGGGATCAC 969  
 Qy 781 TACTTGTGTGATGTGACATGGGCAACCAAGAAAGAAATTTCTTTCAGTGGCTCAGGAGG 840  
 Db 970 TACTTGTGTGATGTGACATGGGCAACCAAGAAAGAAATTTCTTTCAGTGGCTCAGGAGG 1029  
 Qy 841 ATTCAAGACTATTCCGTCAATGGATATTTGTGACTATGATGAATCCAGTGGAAAGTGAAC 900  
 Db 1030 ATTCAAGACTATTCCGTCAATGGATATTTGTGACTATGATGAATCCAGTGGAAAGTGAAC 1089  
 Qy 901 TGCTTAGTGGCAGCGCAACATTTGAATAGTACTACTGCTGGTGGTGGAAAGATTTAGG 960  
 Db 1090 TGCTTAGTGGCAGCGCAACATTTGAATAGTACTACTGCTGGTGGTGGAAAGATTTAGG 1149  
 Qy 961 CCTTCAGAACCTCAATTTTACCTTTGATGGTAAATAGCTTCTCAAGATCATCAGCAATGAA 1020  
 Db 1150 CCTTCAGAACCTCAATTTTACCTTTGATGGTAAATAGCTTCTCAAGATCATCAGCAATGAA 1209  
 Qy 1021 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATAAAGAAAGACTGCAATTTATTACA 1080  
 Db 1210 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATAAAGAAAGACTGCAATTTATTACA 1269  
 Qy 1081 AAAGGCACCTGGGAAGTCAATCGGATAGAGCTCTAACAGTGAATTTATCTATCTACATT 1140  
 Db 1270 AAAGGCACCTGGGAAGTCAATCGGATAGAGCTCTAACAGTGAATTTATCTATCTACATT 1329  
 Qy 1141 AGTAATGAATATAAGGAATCCAGGAGGAAGGAATCTTTATAAAATCCAACTTATTGAC 1200  
 Db 1330 AGTAATGAATATAAGGAATCCAGGAGGAAGGAATCTTTATAAAATCCAACTTATTGAC 1389  
 Qy 1201 TATCAAAAGTGACATGCTCAGTTGTGAGTGAATCCGGAAGAGTGTGAGTACTATTCT 1260  
 Db 1390 TATCAAAAGTGACATGCTCAGTTGTGAGTGAATCCGGAAGAGTGTGAGTACTATTCT 1449  
 Qy 1261 GTGTCTTCACTGAAGGCGCAAGTATATCAGCTGAGATGTTCCGGTCCCTGGTCTGCC 1320  
 Db 1450 GTGTCTTCACTGAAGGCGCAAGTATATCAGCTGAGATGTTCCGGTCCCTGGTCTGCC 1509  
 Qy 1321 CTCTATCTCTACACAGCAGCGTGAATGAATAAGGGCTGAGAGTCCCTGGGAAGACAATTTCA 1380  
 Db 1510 CTCTATCTCTACACAGCAGCGTGAATGAATAAGGGCTGAGAGTCCCTGGGAAGACAATTTCA 1569  
 Qy 1381 GCTTTGGATAAAATGCTGCAAGATGTCCAGATGCCCTCCAAATACTGGACTTCATTATT 1440  
 Db 1570 GCTTTGGATAAAATGCTGCAAGATGTCCAGATGCCCTCCAAATACTGGACTTCATTATT 1629

Qy 1441 TTGAATGAAAACAAATTTTGGTATCAGATGATCTTGCTCTCTCAATTTTGATAAAATCCAAAG 1500  
 Db 1630 TTGAATGAAAACAAATTTTGGTATCAGATGATCTTGCTCTCTCAATTTTGATAAAATCCAAAG 1689  
 Qy 1501 AAATATCTCTCTACTATTTAGATGTTGATGAGGCCCATGTAGTCAAAAGACAGACTGTC 1560  
 Db 1590 AAATATCTCTCTACTATTTAGATGTTGATGAGGCCCATGTAGTCAAAAGACAGACTGTC 1749  
 Qy 1561 TTCAAGCTGAACTGGGCCACTTACCTTGCAGACACAGAAACATATTATGATAGTACTGCTTT 1620  
 Db 1750 TTCAAGCTGAACTGGGCCACTTACCTTGCAGACACAGAAACATATTATGATAGTACTGCTTT 1809  
 Qy 1621 GATGGCAGAGGAAGTGGTTACCAAGGAGATAAGATCATGATGCAATCAACAGAAAGACTG 1680  
 Db 1810 GATGGCAGAGGAAGTGGTTACCAAGGAGATAAGATCATGATGCAATCAACAGAAAGACTG 1869  
 Qy 1681 GGAACTTTGAAAGTGAAGATCAAAATTTGAAGCAGCAGACAAATTTTCAAAAATGGGATTT 1740  
 Db 1870 GGAACTTTGAAAGTGAAGATCAAAATTTGAAGCAGCAGCAGACAAATTTTCAAAAATGGGATTT 1929  
 Qy 1741 GTGGACAAACAAACGAATTTGGGGCTGGTCAATATGGAGGGTACGTAACCTCAATG 1800  
 Db 1930 GTGGACAAACAAACGAATTTGGGGCTGGTCAATATGGAGGGTACGTAACCTCAATG 1989  
 Qy 1801 GTCTCTGGGATCGGGAAGTGGCGTGTCAAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG 1860  
 Db 1990 GTCTCTGGGATCGGGAAGTGGCGTGTCAAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG 2049  
 Qy 1861 TGGGAGTACTATGACTCAGTGTACAGAAAGTTTACATGGGTCTCCCAACTCCAGAAAGAC 1920  
 Db 2050 TGGGAGTACTATGACTCAGTGTACAGAAAGTTTACATGGGTCTCCCAACTCCAGAAAGAC 2109  
 Qy 1921 AACCTTGACCAATTCAGAAATTCAGAGTCAATGAGCAGAGCTGAAATTTTAAACAAAGTT 1980  
 Db 2110 AACCTTGACCAATTCAGAAATTCAGAGTCAATGAGCAGAGCTGAAATTTTAAACAAAGTT 2169  
 Qy 1981 GAGTACCTCTCTTATTCATGGAACAGCAGATGATAACCTTCACTTTACAGCAGTCAAGTCA 2040  
 Db 2170 GAGTACCTCTCTTATTCATGGAACAGCAGATGATAACCTTCACTTTACAGCAGTCAAGTCA 2229  
 Qy 2041 ATCTCCAAAGCCCTGTGATGTTGGAGTGGATTTCCAGGCAATGTGGTATATGATGAA 2100  
 Db 2230 ATCTCCAAAGCCCTGTGATGTTGGAGTGGATTTCCAGGCAATGTGGTATATGATGAA 2289  
 Qy 2101 GACCATGGAATAGTACAGACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2160  
 Db 2290 GACCATGGAATAGTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2349  
 Qy 2161 ATAAACAAATGTTTCTCTTTACCT 2184  
 Db 2350 ATAAACAAATGTTTCTCTTTACCT 2373

RESULT 6

US-11-041-674-5  
 ; Sequence 5, Application US/11041674  
 ; Publication No. US20050181468A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Vanderbilt University  
 ; APPLICANT: Brown, Nancy J.  
 ; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTI  
 ; TITLE OF INVENTION: INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA  
 ; FILE REFERENCE: Acty Docket No. 1242/48/2/2/2  
 ; CURRENT APPLICATION NUMBER: US/11/041,674  
 ; PRIOR FILING DATE: 2005-01-24  
 ; PRIOR APPLICATION NUMBER: 60/244,524  
 ; PRIOR FILING DATE: 2000-10-31  
 ; PRIOR APPLICATION NUMBER: 10/002,593  
 ; PRIOR FILING DATE: 2001-10-31  
 ; PRIOR APPLICATION NUMBER: 10/423,714  
 ; PRIOR FILING DATE: 2003-4-25  
 ; NUMBER OF SEQ ID NOS: 10





QY 2041 ATCTCAAAGCCCTGGTGCATGTTGGAGTGGATTTCCAGGCAATGGGTATCTGATGAA 2100  
 DB 2230 ATCTCAAAGCCCTGGTGCATGTTGGAGTGGATTTCCAGGCAATGGGTATCTGATGAA 2289  
 QY 2101 GACCATTGATAGCTAGCAGCAGCACACACCAACATATATATACCATGAGCCACTTC 2160  
 DB 2290 GACCATTGATAGCTAGCAGCAGCACACACCAACATATATATATACCATGAGCCACTTC 2349  
 QY 2161 ATAAAACAATGTTCTCTTTTACCT 2184  
 DB 2350 ATAAAACAATGTTCTCTTTTACCT 2373

RESULT 7

US-10-295-027-589  
 ; Sequence 589, Application US/10295027  
 ; Publication No. US20030232350A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Afar, Daniel  
 ; APPLICANT: Aziz, Natasha  
 ; APPLICANT: Ginsberg, Wendy M.  
 ; APPLICANT: Gish, Kurt C.  
 ; APPLICANT: Glynn, Richard  
 ; APPLICANT: Hevezi, Peter A.  
 ; APPLICANT: Mack, David H.  
 ; APPLICANT: Murray, Richard  
 ; APPLICANT: Watson, Susan R.  
 ; APPLICANT: Eos Biotechnology, Inc.  
 ; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
 ; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer  
 ; FILE REFERENCE: 018501-012500US  
 ; CURRENT APPLICATION NUMBER: US/10/295,027  
 ; CURRENT FILING DATE: 2002-11-13  
 ; PRIOR APPLICATION NUMBER: US 09/663,733  
 ; PRIOR FILING DATE: 2000-09-15  
 ; PRIOR APPLICATION NUMBER: US 60/350,666  
 ; PRIOR FILING DATE: 2001-11-13  
 ; PRIOR APPLICATION NUMBER: US 60/335,394  
 ; PRIOR FILING DATE: 2001-11-15  
 ; PRIOR APPLICATION NUMBER: US 60/332,464  
 ; PRIOR FILING DATE: 2001-11-21  
 ; PRIOR APPLICATION NUMBER: US 60/334,393  
 ; PRIOR FILING DATE: 2001-11-29  
 ; PRIOR APPLICATION NUMBER: US 60/340,376  
 ; PRIOR FILING DATE: 2001-12-14  
 ; PRIOR APPLICATION NUMBER: US 60/347,211  
 ; PRIOR FILING DATE: 2002-01-08  
 ; PRIOR APPLICATION NUMBER: US 60/347,349  
 ; PRIOR FILING DATE: 2002-01-10  
 ; PRIOR APPLICATION NUMBER: US 60/355,250  
 ; PRIOR FILING DATE: 2002-02-08  
 ; PRIOR APPLICATION NUMBER: US 60/356,714  
 ; PRIOR FILING DATE: 2002-02-13  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 1386  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 589  
 ; LENGTH: 2301  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-295-027-589

Query Match 99.9%; Score 2182.4; DB 6; Length 2301;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 2183; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 AGTCGCAAACTTACACTCTAACTGATTACCTTAAAAAATACTATAGACTCAAGTTATAC 60  
 DB 115 AGTCGCAAACTTACACTCTAACTGATTACCTTAAAAAATACTATAGACTCAAGTTATAC 174  
 QY 61 TCCTTAGATGGATTGATGATGATATCTCTACAAACAAGAAATAATATATCTTGGTA 120

DB 175 TCCTTAGATGGATTGATGATGATATCTCTACAAACAAGAAATAATATCTTGGTA 234  
 QY 121 TTCAATGCTGAATATGGAAACAGCTCAGTTTTCTTGGAGAAACAGTACATTTGATCAGTTT 180  
 DB 235 TTCAATGCTGAATATGGAAACAGCTCAGTTTTCTTGGAGAAACAGTACATTTGATCAGTTT 294  
 QY 181 GGACATTTCTAATGATGATTTCAATATCTCTGATGGGAGTTTATTTCTCTTGAATAC 240  
 DB 295 GGACATTTCTAATGATGATTTCAATATCTCTGATGGGAGTTTATTTCTCTTGAATAC 354  
 QY 241 AACTACGTGAAGCAATGGAGGCATTTCTTACACAGCTTCATATGATGATTTATGATTTAAAT 300  
 DB 355 AACTACGTGAAGCAATGGAGGCATTTCTTACACAGCTTCATATGATGATTTATGATTTAAAT 414  
 QY 301 AAAAGGCAGCTGATTTACAGAAAGAGAGGATTTCAAACAACACACAGTGGGTCACTGGTCA 360  
 DB 415 AAAAGGCAGCTGATTTACAGAAAGAGAGGATTTCAAACAACACACAGTGGGTCACTGGTCA 474  
 QY 361 CCAGTGGGTCAATAATTTGGCATATGTTTGGAAACAATGACATTTATGTTAAATTTGAACCA 420  
 DB 475 CCAGTGGGTCAATAATTTGGCATATGTTTGGAAACAATGACATTTATGTTAAATTTGAACCA 534  
 QY 421 AATTTACCAAGTTACAGAAATCAGATGGAACGGGGAAGGATATATATATATGGAATA 480  
 DB 535 AATTTACCAAGTTACAGAAATCAGATGGAACGGGGAAGGATATATATATATGGAATA 594  
 QY 481 ACTGACTGGGTATTAAGAGAGAGGATTTCTTCACTGCTACTCTGCTCTGCTGGTGGTCTCCA 540  
 DB 595 ACTGACTGGGTATTAAGAGAGAGGATTTCTTCACTGCTACTCTGCTCTGCTGGTGGTCTCCA 654  
 QY 541 AACGGCACTTTTATAGCATATGCCCAATTTAAACGACACAGAAAGTCCCACTTTATGGAATAC 600  
 DB 655 AACGGCACTTTTATAGCATATGCCCAATTTAAACGACACAGAAAGTCCCACTTTATGGAATAC 714  
 QY 601 TCCTTCTACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
 DB 715 TCCTTCTACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 774  
 QY 661 GCAGGAGCTGTAATCCAACTGTAAGTTCTTGTGTGTAATAACAGACTCTCTCAGCTCA 720  
 DB 775 GCAGGAGCTGTAATCCAACTGTAAGTTCTTGTGTGTAATAACAGACTCTCTCAGCTCA 834  
 QY 721 GTCAACCAATGCAACTTCCATACAAACTCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780  
 DB 835 GTCAACCAATGCAACTTCCATACAAACTCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 894  
 QY 781 TACTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
 DB 895 TACTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 954  
 QY 841 ATTACAGAACTATTCCGGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
 DB 955 ATTACAGAACTATTCCGGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1014  
 QY 901 TGCTTAGTGCAACGCGCAACATTTGAATGAGTACTACTGCTGGGTGGGAAGTTTATAGG 960  
 DB 1015 TGCTTAGTGCAACGCGCAACATTTGAATGAGTACTACTGCTGGGTGGGAAGTTTATAGG 1074  
 QY 961 CCTTCAGAACTCAATTTTACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
 DB 1075 CCTTCAGAACTCAATTTTACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1134  
 QY 1021 GAAGGTTACAGACACATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
 DB 1135 GAAGGTTACAGACACATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1194  
 QY 1081 AAAGGCACTTGGGAAGTCACTGGGATAGAGCTCTTAAACAGTGAATATCTATCTACTACATT 1140  
 DB 1195 AAAGGCACTTGGGAAGTCACTGGGATAGAGCTCTTAAACAGTGAATATCTATCTACTACATT 1254  
 QY 1141 AGTAATGATATTAAGGAATGCCAGGAGGAGGATCTTTTATAAAATCCAACTTATTTGAC 1200  
 DB 1255 AGTAATGATATTAAGGAATGCCAGGAGGAGGATCTTTTATAAAATCCAACTTATTTGAC 1314

[illegible]

```

; Publication No. US20050074805A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Kochan, Jarema Peter
; APPLICANT: Martin, Mitchell Lee
; APPLICANT: Robinski, James Andrew
;
; TITLE OF INVENTION: Specific Markers for Diabetes
; FILE REFERENCE: 21270US1
; CURRENT APPLICATION NUMBER: US/10/952,459
; CURRENT FILING DATE: 2004-09-28
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 17
; LENGTH: 3445
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-952-459-17

```

Query Match 99.9%; Score 2182.4; DB 9; Length 3445;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2183; Conservative 0; Mismatches 1; Indels 0;

QY	1	AGTCGCAAACTTACACTCTAACTGATTACTTAAAAAATACTTATAGACTGAGGTATAC	60
DB	225	AGTCGCAAACTTACACTCTAACTGATTACTTAAAAAATACTTATAGACTGAGGTATAC	284
QY	61	TCCTTTAAGATGGATTTTCAGATCATGAATATCTCTACAAAACAAGAAAAATAATATCTTGGTA	120
DB	285	TCCTTTAAGATGGATTTTCAGATCATGAATATCTCTACAAAACAAGAAAAATAATATCTTGGTA	344
QY	121	TTCAATGCTGAATATGGAAAACAGCTCAGTCTTTCTTGGAGAAACAGTACATTTGATCAGTTT	180
DB	345	TTCAATGCTGAATATGGAAAACAGCTCAGTCTTTCTTGGAGAAACAGTACATTTGATCAGTTT	404
QY	181	GGACATCTTATCAATGATTTATTCAAATATCTCTGATGGCGAGTTTATTTCTCTTGAATAC	240
DB	405	GGACATCTTATCAATGATTTATTCAAATATCTCTGATGGCGAGTTTATTTCTCTTGAATAC	464
QY	241	AACTACGTGAAGCAATGGAGSCATTTCTTACACAGCTTCATATGACATTTATGATTTAAAT	300
DB	465	AACTACGTGAAGCAATGGAGSCATTTCTTACACAGCTTCATATGACATTTATGATTTAAAT	524
QY	301	AAAAAGGCAGCTGATTACAGAAAGAGAGGANTTCCAAACAAACACACAGTGGGTACATGGTCA	360
DB	525	AAAAAGGCAGCTGATTACAGAAAGAGAGGANTTCCAAACAAACACACAGTGGGTACATGGTCA	584
QY	361	CCAGTGGGTCAATAATTGGCATAATGTTTGGAAACAATGACATTTATGTTAAAAATTGAAACA	420
DB	585	CCAGTGGGTCAATAATTGGCATAATGTTTGGAAACAATGACATTTATGTTAAAAATTGAAACA	644
QY	421	AAATTTACCAGTTACAGAAACACATGGACGGGGAAGAGATATAATATATAATGGAATA	480
DB	645	AAATTTACCAGTTACAGAAACACATGGACGGGGAAGAGATATAATATATAATGGAATA	704
QY	481	ACTGACTGGGTTTATGAAGAGGAAGTCTTTCAGTGCCTACTCTGCTCTGGTGGTCTCCA	540
DB	705	ACTGACTGGGTTTATGAAGAGGAAGTCTTTCAGTGCCTACTCTGCTCTGGTGGTCTCCA	764
QY	541	AACGGCACTTTTTTATGACATATGCCCAATAATTAAACGACACAGAAAGTCCCACTTATGGAATAC	600
DB	765	AACGGCACTTTTTTATGACATATGCCCAATAATTAAACGACACAGAAAGTCCCACTTATGGAATAC	824
QY	601	TCCTTTCTACTCTGATGAGTCACTCGAGTACCCAAAGACTGTACGGGTTCCATATCCAAAG	660
DB	825	TCCTTTCTACTCTGATGAGTCACTCGAGTACCCAAAGACTGTACGGGTTCCATATCCAAAG	884
QY	661	GCAGGAGCTGTGAATCCAACTGTAAAGTTCTTTTGTGTAAATAACAGACTCTCTCAGCTCA	720
DB	885	GCAGGAGCTGTGAATCCAACTGTAAAGTTCTTTTGTGTAAATAACAGACTCTCTCAGCTCA	944
QY	721	GTCCACCAATGCAACTTCCATACAAATCACTGCTCTGCTCTTATGTTGATGGGATCATC	780
DB	945	GTCCACCAATGCAACTTCCATACAAATCACTGCTCTGCTCTTATGTTGATGGGATCATC	1000







Query Match									
Best Local Similarity 99.9%; Score 2182.4; DB 9; Length 3445;									
Matches 2183; Conservative 0; Mismatches 1; Indels 0; Gaps 0;									
Qy	1	AGTCGCAAACTTACACCTCTAACTGATTACTTAAAAATACTTATAGAGCTGAAGTTATAC	60	Qy	1021	GAAGGTTACAGACACATTTTGTCTATTTCCAAATAGATAAAAAAGACTGCACATTTATTACA	1080		
Db	225	AGTCGCAAACTTACACCTCTAACTGATTACTTAAAAATACTTATAGAGCTGAAGTTATAC	284	Db	1245	GAAGGTTACAGACACATTTTGTCTATTTCCAAATAGATAAAAAAGACTGCACATTTATTACA	1304		
Qy	61	TCCTTAAGATGGATTTACAGATCATGAATATCTCTACAAAACAGAAAATAATATCTTGGTA	120	Qy	1081	AAAGGCACCTCGGAAGTCATCGGGATAGAACTCTAAACAGTGATTATCTATCTACTACATT	1140		
Db	285	TCCTTAAGATGGATTTACAGATCATGAATATCTCTACAAAACAGAAAATAATATCTTGGTA	344	Db	1305	AAAGGCACCTCGGAAGTCATCGGGATAGAACTCTAAACAGTGATTATCTATCTACTACATT	1364		
Qy	121	TTCAATGCTGAATATGAAAACAGCTCAGTTTTCTTGGAGAAACAGTACATTTGATGAGTTT	180	Qy	1141	AGTAATGAATATAAAGGAATGCCAGGAGGAAGGAATCTTTATAAAATCAAATTTATTGAC	1200		
Db	345	TTCAATGCTGAATATGAAAACAGCTCAGTTTTCTTGGAGAAACAGTACATTTGATGAGTTT	404	Db	1365	AGTAATGAATATAAAGGAATGCCAGGAGGAAGGAATCTTTATAAAATCAAATTTATTGAC	1424		
Qy	181	GGACATTTCTATCAATGATTTATCAATATCTCTGATGGCGAGTTTATTTCTCTTAGAATAC	240	Qy	1201	TATACAAAAGTGACATGCTCAGTTGTGAGCTGAATCCGGAAGGTGTCACTACTATTCT	1260		
Db	405	GGACATTTCTATCAATGATTTATCAATATCTCTGATGGCGAGTTTATTTCTCTTAGAATAC	464	Db	1425	TATACAAAAGTGACATGCTCAGTTGTGAGCTGAATCCGGAAGGTGTCACTACTATTCT	1484		
Qy	241	AACACGTAAGCAATGGAGCAATTCCTACACAGCTTCCATATGACATTTATGATTTAAAT	300	Qy	1261	GTGTCATTTCAGTAAGAGCGGAAGTATTTATCAGCTGAGATGTTCCGGTCTCTGGTCTGCC	1320		
Db	465	AACACGTAAGCAATGGAGCAATTCCTACACAGCTTCCATATGACATTTATGATTTAAAT	524	Db	1485	GTGTCATTTCAGTAAGAGCGCGAAGTATTTATCAGCTGAGATGTTCCGGTCTCTGGTCTGCC	1544		
Qy	301	AAAAGGCAGCTGATTACAGAGAGAGGATTTCCAAAACAAACACAGTGGGTCCATGGTCA	360	Qy	1321	CTCTATCTCTACACAGCAGCGTGAATGATATAAGGGCTGAGAGTCTCTGGAAGACAAATTCA	1380		
Db	525	AAAAGGCAGCTGATTACAGAGAGAGGATTTCCAAAACAAACACAGTGGGTCCATGGTCA	584	Db	1545	CTCTATCTCTACACAGCAGCGTGAATGATATAAGGGCTGAGAGTCTCTGGAAGACAAATTCA	1604		
Qy	361	CCAGTGGTCAATAATTTGGCATATGTTTGGAAACATGACATTTATGTTAAAAATTTGAACCA	420	Qy	1381	GCTTTGGATAAAAATGCTGCAGAAATGTCAGATGCCCTCCAAAAAATCTGGACTTCATTATT	1440		
Db	585	CCAGTGGTCAATAATTTGGCATATGTTTGGAAACATGACATTTATGTTAAAAATTTGAACCA	644	Db	1605	GCTTTGGATAAAAATGCTGCAGAAATGTCAGATGCCCTCCAAAAAATCTGGACTTCATTATT	1664		
Qy	421	AAATTTACCAAGTTACAGAAATCACATGGAAGGGAAGGATATAATATAATTAATGGAATA	480	Qy	1441	TTGAATGAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCAITTTTGAATAAATCCAAAG	1500		
Db	645	AAATTTACCAAGTTACAGAAATCACATGGAAGGGAAGGATATAATATAATTAATGGAATA	704	Db	1665	TTGAATGAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCAITTTTGAATAAATCCAAAG	1724		
Qy	481	ACTGACTGGGTTTATGAAGAGGAAGTCTTCAGTGCCTACTCTGCTCTGTGGTGGTCTCCA	540	Qy	1501	AAATATCTCTACTATTAGATGTGATGTCAGGCCCATGTCAGTCAAAAAGCAGACACTGTC	1560		
Db	705	ACTGACTGGGTTTATGAAGAGGAAGTCTTCAGTGCCTACTCTGCTCTGTGGTGGTCTCCA	764	Db	1725	AAATATCTCTACTATTAGATGTGATGTCAGGCCCATGTCAGTCAAAAAGCAGACACTGTC	1784		
Qy	541	AACGGCACTTTTATGACATATGCCCAATTTAAAGCAACAGAAAGTCCCACTTATTAATATAC	600	Qy	1561	TTCAGACTGAACTGGGCCACTTTACCTTGCACAGCA CAGAAAACATTTATAGTAGCTAGCTTT	1620		
Db	765	AACGGCACTTTTATGACATATGCCCAATTTAAAGCAACAGAAAGTCCCACTTATTAATATAC	824	Db	1785	TTCAGACTGAACTGGGCCACTTTACCTTGCACAGCA CAGAAAACATTTATAGTAGCTAGCTTT	1844		
Qy	601	TCCTTTCTACTCTGATGATGACTCAGTACCAGGAGCTGACGGGTTCCATATCCAAAG	660	Qy	1621	GATGGCAGAGGAAGTGGTTACCAAGGAGATAGATCATGCAATGCAATCAACAGAGAGACTG	1680		
Db	825	TCCTTTCTACTCTGATGATGACTCAGTACCAGGAGCTGACGGGTTCCATATCCAAAG	884	Db	1845	GATGGCAGAGGAAGTGGTTACCAAGGAGATAGATCATGCAATGCAATCAACAGAGAGACTG	1904		
Qy	661	GCAGGAGCTGTGAATCCAACTGTAAAGTTCTTTGTTGTAATAACAGACTCTCTCAGCTCA	720	Qy	1681	GGAACTTTTGAAGTTGAAGATCAAATTTGAAGCAGCCAGACAAATTTTCAAAAATGGGATTT	1740		
Db	885	GCAGGAGCTGTGAATCCAACTGTAAAGTTCTTTGTTGTAATAACAGACTCTCTCAGCTCA	944	Db	1905	GGAACTTTTGAAGTTGAAGATCAAATTTGAAGCAGCCAGACAAATTTTCAAAAATGGGATTT	1964		
Qy	721	GTACCAATGCAACTTCCATACAAATCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG	780	Qy	1741	GTGACACAAACGAAATTTGCAATTTGGGGCTGGTCAATGAGGGGTACGTAAACCTCAATG	1800		
Db	945	GTACCAATGCAACTTCCATACAAATCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG	1004	Db	1965	GTGACACAAACGAAATTTGCAATTTGGGGCTGGTCAATGAGGGGTACGTAAACCTCAATG	2024		
Qy	781	TACTTGTGTGATGTGACATGGGCAACACAGAAAGAAATTTCTTTGCAAGTGGCTCAGGAG	840	Qy	1801	GTCTCTGGATCGGGAGTGGGGTTCAGTGTGGAAATAGCCGTGGGGCTGTATCCCGG	1860		
Db	1005	TACTTGTGTGATGTGACATGGGCAACACAGAAAGAAATTTCTTTGCAAGTGGCTCAGGAG	1064	Db	2025	GTCTCTGGATCGGGAGTGGGGTTCAGTGTGGAAATAGCCGTGGGGCTGTATCCCGG	2084		
Qy	841	ATTTCAGAACTATTTCGGTCAATGGATTTTGTGACTATGATGAATCCAGTGGAGATGGAAC	900	Qy	1861	TGGGAGTACTATGACTCAGTGTACACAGAAAGTTTACATGCGGTCTCCAACTCCAGAAAGAC	1920		
Db	1065	ATTTCAGAACTATTTCGGTCAATGGATTTTGTGACTATGATGAATCCAGTGGAGATGGAAC	1124	Db	2085	TGGGAGTACTATGACTCAGTGTACACAGAAAGTTTACATGCGGTCTCCAACTCCAGAAAGAC	2144		
Qy	901	TGCTTAGTGGCAACATTTGAATAGTACTACTGCTGGGTTGGAAGATTTAGG	960	Qy	1921	AACCTTGACCAATTCAGAAAATTCAGACATGACAGAGCTGAAAAATTTTAAAAAAGTT	1980		
Db	1125	TGCTTAGTGGCAACATTTGAATAGTACTACTGCTGGGTTGGAAGATTTAGG	1184	Db	2145	AACCTTGACCAATTCAGAAAATTCAGACATGACAGAGCTGAAAAATTTTAAAAAAGTT	2204		
Qy	961	CCTTCAGAACTCAATTTTACCTTTAGTGAATAGCTTCTACAGAGATCATCAGCAATGAA	1020	Qy	1981	GAGTACCTCTCTTATTTCAATGGAAACAGCAGATGATAACCTTCACTTACAGAGTCAAGCTCAG	2040		
Db	1185	CCTTCAGAACTCAATTTTACCTTTAGTGAATAGCTTCTTACAAGATCATCAGCAATGAA	1244	Db	2205	GAGTACCTCTCTTATTTCAATGGAAACAGCAGATGATAACCTTCACTTACAGAGTCAAGCTCAG	2264		



Db 2325 GACCATGGATAGTAGCAGCAGCACACCAACATATATATATACCCACATGAGCCACTTC 2384  
Qy 2161 ATAAACAATGTTCTCTTTTACT 2184  
Db 2385 ATAAACAATGTTCTCTTTACT 2408

RESULT 11  
US-10-887-553A-216  
; Sequence 216, Application US/10887553A  
; Publication No. US20050085436A1  
; GENERAL INFORMATION:  
; APPLICANT: Garza, Dan  
; APPLICANT: Li, Hao  
; TITLE OF INVENTION: Method to treat conditions associated  
; with insulin signalling dysregulation  
; FILE REFERENCE: 4-33262  
; CURRENT APPLICATION NUMBER: US/10/887,553A  
; CURRENT FILING DATE: 2004-07-08  
; PRIOR APPLICATION NUMBER: 60/485,883  
; PRIOR FILING DATE: 2003-08-07  
; NUMBER OF SEQ ID NOS: 1208  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 216  
; LENGTH: 3913  
; TYPE: DNA  
; ORGANISM: human  
US-10-887-553A-216

Query Match 99.9%; Score 2182.4; DB 9; Length 3913;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2183; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 AGTCGCAAACTTACACTCTAACTGATTACTTAAATAAATACTTATAGACTGAAGTTATAC 60  
Db 676 AGTCGCAAACTTACACTCTAACTGATTACTTAAATAAATACTTATAGACTGAAGTTATAC 735  
Qy 61 TCCTTAAGATGGATTTGAGATCATGAATATCTCTACAAAACAAGAAAATATATCTTTGGTA 120  
Db 736 TCCTTAAGATGGATTTGAGATCATGAATATCTCTACAAAACAAGAAAATATATCTTTGGTA 795  
Qy 121 TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGGAACAGATACATTTGATGAGTTT 180  
Db 796 TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGGAACAGATACATTTGATGAGTTT 855  
Qy 181 GGACATTTCTATCAATGATTATCAATATCTCTGATGGGAGTTTATCTCTTTAGAAATAC 240  
Db 856 GGACATTTCTATCAATGATTATCAATATCTCTGATGGGAGTTTATCTCTTTAGAAATAC 915  
Qy 241 AACTAGCTGAAGCAATGGAGGCAATTCCTACACAGCTTTCATATGACATTTATGATTTAAAT 300  
Db 916 AACTAGCTGAAGCAATGGAGGCAATTCCTACACAGCTTTCATATGACATTTATGATTTAAAT 975  
Qy 301 AAAAGCAGCTGATTACAGAGAGAGATTCCAAACAACACACAGTGGGTCAATGGTCA 360  
Db 976 AAAAGCAGCTGATTACAGAGAGAGATTCCAAACAACACACAGTGGGTCAATGGTCA 1035  
Qy 361 CCAGTGGGTGATATAAATGGCATATGTTGGAAACAATGACATTTATGTTAAATTTGAACCA 420  
Db 1036 CCAGTGGGTGATATAAATGGCATATGTTGGAAACAATGACATTTATGTTAAATTTGAACCA 1095  
Qy 421 AATTACCAAGTTTACAGATTCATGGACGGGGAAGAGATATAATATATATATGATGAATA 480  
Db 1096 AATTACCAAGTTTACAGATTCATGGACGGGGAAGAGATATAATATATATATGATGAATA 1155  
Qy 481 ACTGACTGGGTTTATCAAGGGAAGTCTTCAGTGCCTACTCTGCTCTGTTGGTCTTCCA 540  
Db 1156 ACTGACTGGGTTTATCAAGGGAAGTCTTCAGTGCCTACTCTGCTCTGTTGGTCTTCCA 1215  
Qy 541 AACGGCATTCTTTTACCATATGCCCAATTTAAGCACAAGAGTCCCACTTATTTGAATAC 600  
Db 1216 AACGGCATTCTTTTACCATATGCCCAATTTAAGCACAAGAGTCCCACTTATTTGAATAC 1275

Qy 601 TCCTTTACTCTGATGAGTCACTGCAGTACCCAAAGACTGTACGGGTTCCATATCCAAAG 660  
Db 1276 TCCTTTACTCTGATGAGTCACTGCAGTACCCAAAGACTGTACGGGTTCCATATCCAAAG 1335  
Qy 661 GCAGGAGCTGTGAATCCAACTGTAAGTTCTTTTGTGTAATAACAGACTCTCTCAGCTCA 720  
Db 1336 GCAGGAGCTGTGAATCCAACTGTAAGTTCTTTTGTGTAATAACAGACTCTCTCAGCTCA 1395  
Qy 721 GTCAACAATGCAACTTCCATACAAATCACTGCTCTCTGCTTCTATGTTGATAGGGATCAC 780  
Db 1396 GTCAACAATGCAACTTCCATACAAATCACTGCTCTCTGCTTCTATGTTGATAGGGATCAC 1455  
Qy 781 TACTTGTGATGTGACATGGGCAACAACAAGAAAGAAATTTCTTCAGTGGCTCAGGAGG 840  
Db 1456 TACTTGTGATGTGACATGGGCAACAACAAGAAAGAAATTTCTTCAGTGGCTCAGGAGG 1515  
Qy 841 ATTCAAGAACTATTTCGGTCAATGGATATTGTGCACTATGATGAATCCAGTGGAAAGATGGAAC 900  
Db 1516 ATTCAAGAACTATTTCGGTCAATGGATATTGTGCACTATGATGAATCCAGTGGAAAGATGGAAC 1575  
Qy 901 TGCTTAGTGGCACGGCAACAATTCGAATGAGTACTACTGGCTGGGTTGGGAAGATTTAGG 960  
Db 1576 TGCTTAGTGGCACGGCAACAATTCGAATGAGTACTACTGGCTGGGTTGGGAAGATTTAGG 1635  
Qy 961 CTTTCAGAACCTCATTTTACCCCTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
Db 1636 CTTTCAGAACCTCATTTTACCCCTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1695  
Qy 1021 GAAAGTTACAGACACATTTTGCTATTTCACAAATAGATAAAAAAGACTGCACATTTATTACA 1080  
Db 1696 GAAAGTTACAGACACATTTTGCTATTTCACAAATAGATAAAAAAGACTGCACATTTATTACA 1755  
Qy 1081 AAAGCACTCTGGGAAGTCAATCGGATAGAGTCTTAACAGTGAATATCTATACATTT 1140  
Db 1756 AAAGCACTCTGGGAAGTCAATCGGATAGAGTCTTAACAGTGAATATCTATACATTT 1815  
Qy 1141 AGTAATGATATTAAGGAATGCCAGGAGGAAGGATCTTTATAAATCCAACTTTATTGAC 1200  
Db 1816 AGTAATGATATTAAGGAATGCCAGGAGGAAGGATCTTTATAAATCCAACTTTATTGAC 1875  
Qy 1201 TATACAAAGTGACATGCTCAGTTGTGAGCTGAAATCCGGAAAGGTGTCAGTACTATTCT 1260  
Db 1876 TATACAAAGTGACATGCTCAGTTGTGAGCTGAAATCCGGAAAGGTGTCAGTACTATTCT 1935  
Qy 1261 GTGTCATTCAGTAAAGAGGCGAAGTATTATCAGCTGAGATGTTCCGGTCTCGTGGCCC 1320  
Db 1936 GTGTCATTCAGTAAAGAGGCGAAGTATTATCAGCTGAGATGTTCCGGTCTCGTGGCCC 1995  
Qy 1321 CTCATATCTTACACAGCAGCGTGAATGATTAAGGCGTGAGAGTCTCTGGAAGACAATTCA 1380  
Db 1996 CTCATATCTTACACAGCAGCGTGAATGATTAAGGCGTGAGAGTCTCTGGAAGACAATTCA 2055  
Qy 1381 GCTTTGGATATAAATCTCGAGAAATCTCCAGATGCGCTCCAAAAAACTGGACTTCATTATT 1440  
Db 2056 GCTTTGGATATAAATCTCGAGAAATCTCCAGATGCGCTCCAAAAAACTGGACTTCATTATT 2115  
Qy 1441 TTGAATGAAAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGTATATAATCCAAG 1500  
Db 2116 TTGAATGAAAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGTATATAATCCAAG 2175  
Qy 1501 AAATATCTCTACTATTAGATGATGTATGCAGGCCCATGTAGTCAAAAAAGCAGACACTGTC 1560  
Db 2176 AAATATCTCTACTATTAGATGATGTATGCAGGCCCATGTAGTCAAAAAAGCAGACACTGTC 2235  
Qy 1561 TTCAAGCTGAACTGGGCCACTTTTACCTTGAAGCACAAGAAAAATTTATAGTAGTACTGCTTT 1620  
Db 2236 TTCAAGCTGAACTGGGCCACTTTTACCTTGAAGCACAAGAAAAATTTATAGTAGTACTGCTTT 2295  
Qy 1621 GATGGAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGCAATGCAATCAACAGAGACTG 1680  
Db 2296 GATGGAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGCAATGCAATCAACAGAGACTG 2355

[illegible]

**RESULT 12**

US-10-641-643-1026  
; Sequence 1026, Application US/10641643  
; Publication No. US20040077003A1  
; GENERAL INFORMATION:  
; APPLICANT: Cocks, Benjamin G.  
; Susan G. Stuart  
; Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL  
; GENE EXPRESSION  
; NUMBER OF SEQUENCES: 1508  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/641,643  
; FILING DATE: 14-Aug-2003  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: <Unknown>  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0001 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555

QY 841 ATTGAGAACTATTTCGGTCATGGATATTGTGACTATGATGAATCCAGTGGAGATGGAAC 900  
DB 965 ATTCAGAACTATTTCGGTCATGGATATTGTGACTATGATGAATCCAGTGGAGATGGAAC 1024  
QY 901 TGCCTTAGTGCGACGGCAACACATGTGAAATGAGTACTTGGCTGGGTGGAGATTTAGG 960  
DB 1025 TGCCTTAGTGCGACGGCAACACATGTGAAATGAGTACTTGGCTGGGTGGAGATTTAGG 1084  
QY 961 CCTTCAGAACCTCATTTTACCCCTTGATGGTAAATAGCTTCTTACAGATCATCAGCATGAA 1020  
DB 1085 CCTTCAGAACCTCATTTTACCCCTTGATGGTAAATAGCTTCTTACAGATCATCAGCATGAA 1144  
QY 1021 GAAGGTTTACAGACACATTTGCTATTTCCAAATAGATAAAAAAGACTGCACATTTATTACA 1080  
DB 1145 GAAGGTTTACAGACACATTTGCTATTTCCAAATAGATAAAAAAGACTGCACATTTATTACA 1204  
QY 1081 AAGGCACTGGGAAGTCACTGGGATAGAAAGTCTTAACCAAGTATTAATCTATCTACTACATT 1140  
DB 1205 AAGGCACTGGGAAGTCACTGGGATAGAAAGTCTTAACCAAGTATTAATCTATCTACTACATT 1264  
QY 1141 AGTAATGNAATATAAGGAATGCCAGGAGGAGGAATCTTTATAAAATCCAACTTATTGAC 1200  
DB 1265 AGTAATGNAATATAAGGAATGCCAGGAGGAGGAATCTTTATAAAATCCAACTTATTGAC 1324  
QY 1201 TATACAAAAGTGACATGCCCTCAGTGTGAGCTCAATCCGAAAGGTGTCAGTACTATTCT 1260  
DB 1325 TATACAAAAGTGACATGCCCTCAGTGTGAGCTCAATCCGAAAGGTGTCAGTACTATTCT 1384  
QY 1261 GTGTCAATTCACTAAAGAGCGGAAGTATTATCAGCTGAGATGTTCCGGTCTGGTCTGCC 1320  
DB 1385 GTGTCAATTCACTAAAGAGCGGAAGTATTATCAGCTGAGATGTTCCGGTCTGGTCTGCC 1444  
QY 1321 CTCTATCTACTACAGCAGCGTGAATGATTAAGGGCTGAGAGTCTCTGGAAGCAATTTCA 1380  
DB 1445 CTCTATCTACTACAGCAGCGTGAATGATTAAGGGCTGAGAGTCTCTGGAAGCAATTTCA 1504  
QY 1381 GCTTTGGATAAATGCTGCAGAAATGTCAGATGCCCTCCAAAAAAGCTGGACTTCATTATT 1440  
DB 1505 GCTTTGGATAAATGCTGCAGAAATGTCAGATGCCCTCCAAAAAAGCTGGACTTCATTATT 1564  
QY 1441 TTGAATGAAACAAAAATTTTGGTATCAGATGATCTTGCCTCCTCATTTTGTATAAATCCAAG 1500  
DB 1565 TTGAATGAAACAAAAATTTTGGTATCAGATGATCTTGCCTCCTCATTTTGTATAAATCCAAG 1624  
QY 1501 AATATCCTCTACTATTAGATGTGATGACGGCCCATGTAGTCAAAAAAGCAGACACTGTC 1560  
DB 1625 AATATCCTCTACTATTAGATGTGATGACGGCCCATGTAGTCAAAAAAGCAGACACTGTC 1684  
QY 1561 TTGAGACTGAACTGGGCCCACTTACCTTGCAGCAGACAGAAAAACATTTATAGTAGCTGTT 1620  
DB 1685 TTGAGACTGAACTGGGCCCACTTACCTTGCAGCAGACAGAAAAACATTTATAGTAGCTGTT 1744  
QY 1621 GATGGCAGGAAGTGGTTACCAAGAGATAAGATCATGCTCAATCAACAGAGACTG 1680  
DB 1745 GATGGCAGGAAGTGGTTACCAAGAGATAAGATCATGCTCAATCAACAGAGACTG 1804  
QY 1681 GGAACATTTGAAGTTGGAAGTCAAAATTTGAAGCCAGACAAATTTTCAAAAATGGGATTT 1740  
DB 1805 GGAACATTTGAAGTTGGAAGTCAAAATTTGAAGCCAGACAAATTTTCAAAAATGGGATTT 1864  
QY 1741 GTGGACAAACAAACGAATTTGCAATTTTGGGGCTGGTCAATATGGAGGGTACCTTAACCTCAATG 1800  
DB 1865 GTGGACAAACAAACGAATTTGCAATTTTGGGGCTGGTCAATATGGAGGGTACCTTAACCTCAATG 1924  
QY 1801 GTCTGGGATCGGGAAAGTGGCGTGTTCAGATGTGGAAATAGCCGTGGCGCTGTATCCCGG 1860  
DB 1925 GTCTGGGATCGGGAAAGTGGCGTGTTCAGATGTGGAAATAGCCGTGGCGCTGTATCCCGG 1984  
QY 1861 TGGGAGTACTATGACTCAGTGTACACAGAAAGTTTACATGGGTCTCCCAATCCAGAGAC 1920  
DB 1985 TGGGAGTACTATGACTCAGTGTACACAGAAAGTTTACATGGGTCTCCCAATCCAGAGAC 2044  
QY 1921 AACCTTGACCATTAACAGAAATTCACAGTCATGAGCAGAGCTGAAAAATTTTAAACAGTT 1980

DB 2045 AACCTTGACCATTAACAGAAATTCACAGTTCATGAGCAGAGCTGAAAAATTTTAAACAGTT 2104  
QY 1981 GAGTACCTCCTTATTATTCATGGAAACAGCAGATGATAGTTCACCTTTCAGCAGTCAGCTCAG 2040  
DB 2105 GAGTACCTCCTTATTATTCATGGAAACAGCAGATGATAGTTCACCTTTCAGCAGTCAGCTCAG 2164  
QY 2041 ATCTCCAAAGCCCTGGTCGATGTTGGAGTGGATTTCCAGGCAATGTGGTATATCTGATGAA 2100  
DB 2165 ATCTCCAAAGCCCTGGTCGATGTTGGAGTGGATTTCCAGGCAATGTGGTATATCTGATGAA 2224  
QY 2101 GACCATGGAATAGCTAGCAGCAGCAGCACCAACATATATATATACCCACATGAGCCACTTC 2160  
DB 2225 GACCATGGAATAGCTAGCAGCAGCAGCACCAACATATATATATATACCCACATGAGCCACTTC 2284  
QY 2161 ATAAAAACAATGTTTCTCTTTTACCT 2184  
DB 2285 ATAAAAACAATGTTTCTCTTTTACCT 2308  
  
RESULT 13  
US-10-631-467-969  
; Sequence 969, Application US/10631467  
; Publication No. US20050208496A1  
; GENERAL INFORMATION:  
; APPLICANT: Genox Research Inc.  
; TITLE OF INVENTION: Method for testing for broncheal asthma, or chronic obstructive f  
; FILE REFERENCE: 3462.1005-000  
; CURRENT APPLICATION NUMBER: US/10/631,467  
; CURRENT FILING DATE: 2003-07-31  
; PRIOR APPLICATION NUMBER: JP 2003-077212  
; PRIOR FILING DATE: 2003-03-20  
; PRIOR APPLICATION NUMBER: JP 2002-229312  
; PRIOR FILING DATE: 2002-08-06  
; NUMBER OF SEQ ID NOS: 2086  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 969  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-10-631-467-969  
  
Query Match 76.8%; Score 1677.6; DB 9; Length 3316;  
Best Local Similarity 86.0%; Pred. No. 0;  
Matches 1876; Conservative 0; Mismatches 294; Indels 12; Gaps 1;  
  
QY 1 AGTCGCAAAAACCTTACACTCACTGATTACTTAAAAAATACTTATAGACTGAAGTTATAC 60  
DB 226 AGCCGCGAGAACGTTATTCAGTACCTGACTATTTAAAGAGTACCTTTCCGGTCAAGTCCCTAC 285  
QY 61 TCCTTAAGATGGATTTTCAGATCATGAATATCTCTCAAAACAAGAAAAATAATCTTGGTA 120  
DB 286 TCTTTGTGGTGGTTTCAGACTTTTGAATACCTCTACAAACAGAGAACAAATATCTTGGCTG 345  
QY 121 TTCATGCTGATATGGAACAGCTCAGTCTTTCTGGAGAACAGTACATTTTCATGAGTTT 180  
DB 346 CTCATGCTGAACATGGAAACAGTCCCATTTCTTGGAGAACAGTACCTTTGAAAGCTTT 405  
QY 181 GGACATTTCTCAATGATTATTCTCAATATCTCTGATGGCAGTTTATTCTTTAGAAATAC 240  
DB 406 GGA-----TATCATTTCACTGCTGCACTGACCCAGCTGTTTCTCTTGGNAATAC 453  
QY 241 AACTACGTGAAGCAATGGAGCATTCCTACACAGCTTCATATGACATTTATGATTTAAAT 300  
DB 454 AACTACGTGAAGCAATGGAGCATTCCTACACAGCTTCATATGACATTTATGATGTAAT 513  
QY 301 AAAAGGCACTGATTACAGAGAGAGGATTCGAACACACACACAGTGGGTGCATGGTCA 360  
DB 514 AAAAGCACTGATCACAGAGAGAGGATTCGAATAATATACACAGTGGATTCATGGTCA 573  
QY 361 CCAGTGGGTCTATAAATTTGGCATATGTTTGGAAACATGACATTTATGTTAAATTTGAACCA 420



```

; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1570
; LENGTH: 4835
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_012789
US-09-917-800A-1570

Query Match          75.8%; Score 1656.2; DB 3; Length 4835;
Best Local Similarity 85.3%; Pred. No. 0;
Matches 1863; Conservative 0; Mismatches 313; Indels 9; Gaps 1;

QY      7  AAAAATTACACTCTAACTGATTAATCTCTAATAAAGAAATTAATATCTTGGTATTCAT 126
D      203 AGAATTACACTAGCTAGCTATATTAAGAAATACCTTTGGGTCAAGTCTCTACTCTTG 262

QY      67  AGATGGATTCAGATCATGAATATCTCTAATAAAGAAATTAATATCTTGGTATTCAT 126
D      263 CGTGGGTTTCAGATCTCTGAATACCTCTAAGCAAGAAATCAATATCTTGTATTCAT 322

QY      127 GCTGAATATGGAACACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTGGACAT 186
D      323 GCTGAAACAGGGAACAGCTCCATTTCTTGGAGAACAGTACCTTTTGAGATCTTTGGAGAT 382

QY      187 TCTATCAATGATTATTCATATCTCTGATGGCAGTTTATCTCTTGAATACAACTAC 246
D      383 TCTATNAAGTATTTAGTGTGACCCGACAGACTGTTGTTCTCTTAGAATACAAATAT 442

QY      247 GTGAAGCAATGGAGGCAATCTCTACACAGCTTCATATGACATTTATGATTTAAATAAAGG 306
D      443 GTGAGCAATGGAGACACTCTACACGGCTTCATACAGTATTTATGACTTGAATAAAGA 502

QY      307 CAGCTGATTCAGAGAGAGATTCAGAACACACACAGTGGGTGATGCTCAGTCCAGTG 366
D      503 CAGCTGATTCAGAGAGAGATTCAGAACATATATACACAGTGATCAATGGTCAACAGAA 562

QY      367 GGTCAATAATTCGCATATGTTGGAAATGATGATTTATGTTAAATTCGAACCAATTTA 426
D      563 GGTCAATAATTCGCATATGTTGGAAATGATGATTTATGTTAAATTCGAACCAATTTG 622

QY      427 CCAAGTTACAGAAATCACTGGACGGGAAAGAGATATATATATATGAAATTAACCTGAC 486
D      623 CCAAGTTACAGAAATCACTGGACGGGAAAGAGATATATATATGAAATTAACCTGAC 682

QY      487 TGGGTTTATGAGAGAGAGTCTTCAGTGCTACTCTGCTCTGTTGGTGGTCTCAACGGC 546
D      683 TGGGTTTATGAGAGAGAGTCTTCAGTGCTACTCTGCTCTGTTGGTGGTCTCAACGGC 742

QY      547 ACTTTTATGACATATGCCCAATTTAACGACACAGAGTCCCACTTTATGAAATCTCTTC 606
D      743 ACTTTTATGACATATGCCCAATTTAACGACACAGAGTCCCACTTTATGAAATCTCTTC 802

QY      607 TACTCTGATGAGTCACTGAGTACCCCAAGAGCTGTACGGGTTCCATATCCAAAGGAGGA 666
D      803 TACTCTGATGAGTCACTGAGTACCCCAAGAGCTGTGATCCGTTACCCAAAGGAGGA 862

QY      667 GCTGGAATCCAACTGTAAGTTCTTTGTTGTAATACAGACTCTCTCAGCTCAGTCAAC 726
D      863 GCTGGAATCCAACTGTAAGTTCTTTGTTGTAATACAGACTCTCTCAGCTCAGTCAAC 922

QY      727 AATGCAACTTCCATACAAATCACTGCTCTGCTCTCTATGTTGATAGGGATCACTACTTG 786

```

```

D      923 ACTACGATTCCTCATGCAAAATCACCGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 982
QY      787 TGTGATGTCATGCGGCAACACAGAAAGAAATTTCTTTCAGTGGCTCAGGAGGATTCAG 846
D      983 TGTGACGTGGCTGGGTTTCAGAAAGACAGAAATCTCTGTCAGTGGCTCAGGAGGATTCAG 1042
QY      847 AACTATTCGGTCATGATATTTTGTGACTATGATGAAATCCAGTGGAGATGGAACCTGCTTA 906
D      1043 AACTATTCGGTCATGATGCGGATCTGCGACTATGATGAAACACACCTAGTATGGAACCTGCTTA 1102
QY      907 GTGGCAGCGCAACACATTTGAAATGAGTACTACTGCTGGGTTTGGAGAAATTTAGGCTTCA 966
D      1103 ACGACGAGGAGCAATTTGAAACGAGTCCACAGGCTGGTGGGAAATTTAGGCTGCA 1162
QY      967 GAACCTCAATTTTACCTTGTGATGATGATGCTTCTCAAGATCATCAGCAATGAGAGGT 1026
D      1163 GAACCCCACTTCACTCCGACGAGGAGCTCTCTATAAATCTGTCAGTGAACAAAGATGCG 1222
QY      1027 TACAGACATTTGCTATTTTCCAAATAGATA-----AAAAAGACTGCACATTTAT 1077
D      1223 TACAACACATCTGCCAGTTCAGAAAGATAGGAAACCGGAACAGGTCTGTACATTTAT 1282
QY      1078 ACAAAGGCACTGGGAAAGTCACTCGGATAGAGCTCTAACCCAGTGTATCTATCTATAC 1137
D      1283 ACANAAGGAGCTGGGAAAGTCAATAGTATCGAAGCTCTGACCCAGCGATTTATCTGTACTAC 1342
QY      1138 ATTAGTAATGAATATAAAGGAATGCCAGGAGGAAGAAATCTTTATAAATTCACCTTAT 1197
D      1343 ATTAGTAATGAATATAAAGAAATGCCAGGAGGAAGAAATCTTTATAAATTCAGCTTACT 1402
QY      1198 GACTATACAAAGTCACTGCTCAGTTGTGAGCTGAATCCGAAAGGTGTCAGTACTAT 1257
D      1403 GACCAACAATAGAGAGTCCCTTAGTTGTGACCTGAAATCCGAAAGATGCGCAGTATATAC 1462
QY      1258 TCTGTGTCAATTCAGTAAAGAGCGGAAGTATTTATCAAGTGTGAGATGTTCCGGTCTCTG 1317
D      1463 TCGGTGTCACTTAGTAAAGAGCAAGTATATCAAGTGGGATCGCGGGCCCTGGTCTG 1522
QY      1318 CCCCTCTATCTCTACAGCAGCGGTGAATGATTAAGGCTCAGAGTCTCTGGAAGCAAT 1377
D      1523 CCCCTCTACACTCTGATCGCAGCACTGATCAAAAAGAGCTGAGAGTCTCTGGAAGCAAT 1582
QY      1378 TCAGCTTTGGGATAAATGCTGAGAAATGTCAGAGTCCCTCCAAAGAACTGGACTTCAT 1437
D      1583 TCTGTTTGGATAAATGCTGCAAGTGTCCAAATGCTTCCAAATTAATTTGACTTCAT 1642
QY      1438 ATTTTGAATGAAACAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTGATAAATCC 1497
D      1643 GTTCTGATGAAACAAAGATTTTGGTATCAATGATCTTACCTCTCTCATTTGATAAATCC 1702
QY      1498 AAGAAATATCTCTATCTATGATGATGATGAGCGCCCATGATGATGATGATGATGATGAT 1557
D      1703 AAGAAATATCTCTATCTATGATGATGATGAGCGCCCATGATGATGATGATGATGATGAT 1762
QY      1558 GTCTTCAGACTCACTGGGCTCCTTACCTTGAACACAGAAATCAATATGATGATGATGATGAT 1617
D      1763 GCCTTCAGACTCACTGGGCTCCTTACCTTGAACACAGAAATCAATATGATGATGATGATGAT 1822
QY      1618 TTTGATGCGAGAGGAGTGGTTTACCAAGGAGATTAAGATCATGATGATGATGATGATGATGAT 1677
D      1823 TTTGATGCGAGAGGAGTGGTTTACCAAGGAGATTAAGATCATGATGATGATGATGATGATGAT 1882
QY      1678 CTGGAAACATTTGAAAGTTGAAGATCAAAATGGAAGCGGAGCAATTTTCAAAAATGGGA 1737
D      1883 CTGGAAACATTTGAAAGTTGAAGATCAAAATGGAAGCGGAGCAATTTTCAAAAATGGGA 1942
QY      1738 TTTTGGGACACAAAGAAATTTGCAATTTGGGCTGCTCATATGAGGGTACGTTAACTCA 1797
D      1943 TTTTGGGACACAAAGAAATTTGCAATTTGGGCTGCTCATATGAGGGTACGTTAACTCA 2002
QY      1798 ATGTGCTCTGGGATCGGGAAGTGGCGTGTTCAGTGTGGAATAGCCGTGGCGCTGTATCC 1857

```

Db 2003 ATGGTCTGGGANTCGGGAAGTGGCGTGTTCAGTGTGGAATAGCCGTGGCGCCCGTGTCA 2062  
Qy CGGTGGGAGTACTATGACTCAGTGTACACAGAACTTTACATGGGTCTCCCACTCCAGAA 1917  
Db 2063 CGGTGGGAGTACTATGACTCAGTGTACACAGAGCGTTACATGGGTCTCCCACTCCAGAG 2122  
Qy 1918 GACAACTTGACCAATTAAGAAATTCACAGTCAATGACAGAGCTGAAATTTTAAACAA 1977  
Db 2123 GACAACTTGACCAATTAAGAAATTCACAGTCAATGACAGAGCTGAAATTTTAAAGCAA 2182  
Qy 1978 GTTCAGTACCTCTCTATTTATGGAACAGACAGATGATACGTTTCACTTTACAGAGTCAGCT 2037  
Db 2183 GTTCAGTACCTCTCTATTTACAGGTACAGAGATGATATGTTCACTTTACAGAGTCAGCT 2242  
Qy 2038 CAGATCTCAAAAGCCCTCGTGCATGTGGAGTGGATTTCCAGGCAATGTGGTATACTGAT 2097  
Db 2243 CAGATCTCAAAAGCCCTCGTGCATGTGGAGTGGATTTCCAGGCAATGTGGTATACGGAC 2302  
Qy 2098 GAAGACCAATGAATAGCTAGCAGCAGCAGCACACCAACATATATATACCAATAGAGCCAC 2157  
Db 2303 GAAGACCAATGGATCGCCAGCAGCAGCAGCTCACAGCACATCTATTCCACATAGGCCAT 2362  
Qy 2158 TTCATAAACAAATGTTCTCTTTAC 2182  
Db 2363 TTCTCCAGCAGTGTCTCTCTTTAC 2387

RESULT 15  
US-10-165-603-5  
; Sequence 5, Application US/10165603  
; Publication No. US20030021792A1  
; GENERAL INFORMATION:  
; APPLICANT: Roben, Paul W.  
; APPLICANT: Stevens, Anthony C.  
; TITLE OF INVENTION: TISSUE-SPECIFIC ENDOTHELIAL MEMBRANE  
; FILE REFERENCE: TPTECH.001A  
; CURRENT APPLICATION NUMBER: US/10/165,603  
; PRIOR FILING DATE: 2002-06-07  
; PRIOR APPLICATION NUMBER: 60/297,021  
; PRIOR FILING DATE: 2001-06-08  
; PRIOR APPLICATION NUMBER: 60/305,117  
; PRIOR FILING DATE: 2001-07-12  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 4835  
; TYPE: DNA  
; ORGANISM: Rattus norvegicus  
US-10-165-603-5

Query Match 75.8%; Score 1656.2; DB 5; Length 4835;  
Best Local Similarity 85.3%; Pred. No. 0;  
Matches 1863; Conservative 0; Mismatches 313; Indels 9; Gaps 1;

Qy 7 AAACTTACACTCTAACTGATTACTTAAATAATCTATAGACTGAAGTTATCTCTTCA 66  
Db 203 AGAATTACACACTAGTCACTATTTAAAGAAATACCTTTCGGGTCAAGTCTCTCTT 262  
Qy 67 AGATGGATTTAGATCATGATATCTCTCAAAACAAGAAATAATATCTTGGTATTCAT 126  
Db 263 CGGTGGGTTCAGATTCCTGAATACCTCTACAAGCAAGAAACAATATCTTCTTATTCAT 322  
Qy 127 GCTCAATATGGAACAGCTAGTTTCTTGGAGACAGTACATTTGATGATTTGGACAT 186  
Db 323 GCTGAACACGGGAACAGCTCCATTTTCTTGGAGAACAGTACCTTTTGAGATCTTTGGAGAT 382  
Qy 187 TCTATCAATGATTTTCAATATCTCTGATGGGAGTTTATCTCTTAGAATACAACTAC 246  
Db 383 TCTATAGTATTTTCACTGTCACTCCGACAGACTGTTCGTCTCTTAGAATACAAATAT 442  
Qy 247 GTGAAGCAATGGAGGCAATTCCTACACAGCTTCATATGACATTTATGATTTAAATAAAGG 306

Db 443 GTGAAGCAATGGAGACACTCTCTACACGGCTTCATACAGTATTTATGACTTGAATAAAGA 502  
Qy 307 CAGCTGATTACAGAAGAGAGGATTTCCAAACAACAACACACAGTGGGTCACTAGTGCACAGTG 366  
Db 503 CAGCTGATCACAGAAGAGAGATTTCCAAATATATACACAGTGGATCATAGTGTCAAGAA 562  
Qy 367 GGTCAATAAATTTGGCATATGTTTGGAAACAATCACAATTTATGTTAAATTTGAACCAATTTA 426  
Db 563 GGTCAATAAATTTGGCATATGTTCTGCAAGAAATGATATTTATGTTAAATTTGAACCAATTTG 622  
Qy 427 CCAAGTTACAGAAATCACATGACCGGGNAAGAAATATAATATATATATGGAATACTGAC 486  
Db 623 CCTAGTCATAGGATCACATCAACAGGAAAGAAATGTAATATTTAACGGAATAAATGAC 682  
Qy 487 TGGGTTTATGAAGAGAGAGTCTTTCAGTGCCTACTCTGCTCTGTGTGTGTCTCTCAACCGC 546  
Db 583 TGGGTTTATGAAGAGAGAAATCTTTCGGTGCCTACTCTCGCACATGTGTGTGTCTCTCAACCGC 742  
Qy 547 ACTTTTTTAGCATATGCCCAATTTAAGGACACAGAGTCCCACTTATTTGAATATCTCTTTC 606  
Db 743 ACTTTTCTAGCTTATGCCCACTTTAAACGACACCGAGTGCCTCTCATTTGAATATCTCTTTC 802  
Qy 607 TACTCTGATGAGTCACTGACAGTACCCAAAGACTGTAACGGGTTCATATCCAAAGCAGGA 666  
Db 803 TACTCTGATGAGTCACTGACAGTACCCCAAGCACTGCTGGATTCGGTACCCAAAGCAGGA 862  
Qy 667 GCTGTGAATCCAACTGTAAAGTCTTTTGTGTAATACAGACTCTCTCAGCTCAGTCACTC 726  
Db 863 GCTGTGAATCCAACTGTAAAGTCTTTTATGTAATACAGACTCTCTCAGCTCAGTCACTC 922  
Qy 727 AATGCAACTTCATACAAATCACTGCTCTGCTCTTCTATGTTGATAGGGGATCACTACTTG 786  
Db 923 ACTACGATTTCCCAATGCAAAATCACCGCTCTCTGATCTGTGACACACAGGGGATCACTACTTG 982  
Qy 787 TGTGATGTGATCGGCAACACAGAAAGAAATTTCTTTCAGTGGCTCAGAGGATTCAG 846  
Db 983 TGTGACGTGGCTGGGTTTCAGAAAGACAGAAATCTCGTTGCACTGGCTCAGGAGGATTCAG 1042  
Qy 847 AACTATTTCGGTCACTGGATATTTGTGACTATGATGAATCCAGTGGAAAGATGGAACCTGCTTA 906  
Db 1043 AACTATTTCGGTCACTGGATATTTGTGACTATGATGAATCCAGTGGAAAGATGGAACCTGCTTA 1102  
Qy 907 GTGCAACGGCAACACATTTGAATCAGTACTACTGCTGGGTGGGTTGGAAGATTTAGGCTTCA 966  
Db 1103 ACGACGCGAGGACATATTTGAAACAGTGCACAGGCTGGTGGCGGAAGATTTAGGCTTCA 1162  
Qy 967 GAACCTCATTTTACCTTGATGTAATAGCTTCTTACAAGATCATCAGCAATGAGAAAGGT 1026  
Db 1163 GAACCCCACTTCACCTCCGACGGAAGCAGCTTCTATAAAATCGTCAGTGACAAAGATGGC 1222  
Qy 1027 TACAGACACATTTGCTATTTCCAAATAGATA- - - - -AAAAAGACTGCACATTTATT 1077  
Db 1223 TACAAACACATCTCCAGTTCAGAAAGATAGGAAACCCGAAACAGTCTGTACATTTATT 1282  
Qy 1078 ACAAAGGCACCTGGGAAGTCATCGGATAGAAGCTCTTAACACAGTGAATTTCTATPACTAC 1137  
Db 1283 ACAAAGGCACCTGGGAAGTCATTTAGTATCGAAGCTCTGACACGAGATTTATCTGTACTAC 1342  
Qy 1138 ATTAGTATATGATATAAGGAATGCCAGGAGGAAGGAATCTTTTATAAAATCCAACTTATT 1197  
Db 1343 ATTAGTATGAAATATAAGGAATGCCAGGAGGAAGAAATCTTTTATAAAATTCAGCTTACT 1402  
Qy 1198 GACTATACAAAGTGACATGCTCAGTTGTGAGTGTGAATCCGGAAGGTGTGAGTACTAT 1257  
Db 1403 GACCACAAATATAAGAGTGCCTTAGTTGTGACCTGAAATCCAGAAAGATGCAATATTAC 1462  
Qy 1258 TCTGTGTCAATTCAGTAAAGAGCGGAAGTATTTATCAGCTGAGATGTTCCGGTCTGGTCTG 1317  
Db 1463 TCGGTGTCACTTAGTAAAGAGCGCAAGTACTATCAGCTGGGATGCCCGGGCCCTGGTCTG 1522  
Qy 1318 CCCCTCTATCTTACACAGCAGCGGTGAATGATAAAGGGCTGAGAGTCTCTGGAAGACAA 1377  
Db 1523 CCCCTCTACACTCTGCACTCGCAGCACTGATCAAAAGAGAGCTGAGAGTCTCTGGAGGACAA 1582





GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2006, 04:36:39 ; Search time 400 Seconds  
(without alignments)  
9705.487 Million cell updates/sec

Title: US-10-659-055-2  
Perfect score: 2184  
Sequence: 1 agtcgcaaaactacactct.....aacaaatgtttcttcttacct 2184

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues  
Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/1 COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5 COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/H COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*  
7: /cgn2\_6/ptodata/1/ina/PP COMB.seq.\*  
8: /cgn2\_6/ptodata/1/ina/RE COMB.seq.\*  
9: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2184	100.0	3407	3	US-10-002-593-5
2	2184	100.0	3407	3	US-09-949-016-275
3	2184	100.0	3407	3	US-10-423-714-5
4	2182.4	99.9	3419	3	US-09-949-016-4579
5	2180.8	99.9	2924	3	US-09-021-655-1026
6	2180.8	99.9	2924	6	PCT-US93-07923-1
7	751	34.4	2815	2	US-08-230-491A-1
8	751	34.4	2815	2	US-08-619-280A-1
9	751	34.4	2815	2	US-08-940-391-1
10	751	34.4	2815	3	US-09-265-606-1
11	669.4	30.7	2736	3	US-09-949-016-2459
12	264.8	12.1	535	3	US-09-389-681-428
13	264.8	12.1	535	3	US-09-620-405B-428
14	264.8	12.1	535	3	US-09-433-826B-428
15	264.8	12.1	535	3	US-09-604-287A-428
16	264.8	12.1	535	3	US-09-834-759-428
17	264.8	12.1	535	3	US-09-590-751A-428
18	264.8	12.1	535	3	US-09-551-621-428
19	264.8	12.1	535	3	US-09-551-621A-428
20	264.8	12.1	535	3	US-10-076-622-428
21	215.8	9.9	2576	3	US-10-363-937-21
22	215.8	9.9	2583	3	US-09-976-674-6
23	215.8	9.9	4541	3	US-09-976-674-42
24	197	9.0	85869	3	US-09-949-016-12017

25	197	9.0	85878	3	US-09-949-016-16321	Sequence 16321, A
26	177.4	8.1	4496	3	US-09-976-674-44	Sequence 44, Appl
c 27	141	6.5	239	3	US-09-392-184-32	Sequence 32, Appl
28	121.8	5.6	1023	3	US-09-270-767-13726	Sequence 13726, A
c 29	109.8	5.0	601	3	US-09-949-016-23690	Sequence 23690, A
c 30	109.8	5.0	601	3	US-09-949-016-163433	Sequence 163433, A
c 31	99	4.5	657	3	US-09-221-017B-646	Sequence 646, App
32	92.6	4.2	4280	3	US-09-079-592-1	Sequence 1, Appl
33	92.6	4.2	4280	3	US-09-462-284-1	Sequence 1, Appl
34	88	4.0	1485	3	US-09-248-796A-5369	Sequence 5369, Ap
c 35	88	4.0	4982	3	US-08-699-103B-1	Sequence 1, Appl
c 36	88	4.0	4982	3	US-09-229-059-1	Sequence 1, Appl
c 37	88	4.0	4982	3	US-09-628-133-1	Sequence 1, Appl
38	81	3.7	547	3	US-09-280-116-249	Sequence 249, App
39	77.8	3.6	2671	3	US-09-976-674-2	Sequence 2, Appl
40	77.8	3.6	2797	3	US-09-976-594-1103	Sequence 1103, Ap
41	77.8	3.6	3120	3	US-10-070-464-2	Sequence 2, Appl
42	77.8	3.6	4309	3	US-09-976-674-14	Sequence 14, Appl
43	77.8	3.6	4685	3	US-09-976-674-22	Sequence 22, Appl
44	77.8	3.6	4685	3	US-09-976-674-12	Sequence 12, Appl
45	76.6	3.5	76848	3	US-09-949-016-14201	Sequence 14201, A

ALIGNMENTS

RESULT 1  
US-10-002-593-5  
; Sequence 5, Application US/10002593  
; Patent No. 6586198  
; GENERAL INFORMATION:  
; APPLICANT: Vanderbilt University  
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERT  
; TITLE OF INVENTION: INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA  
; FILE REFERENCE: Atty Docket No. 6586198 1242/48/2  
; CURRENT APPLICATION NUMBER: US/10/002.593  
; PRIOR FILING DATE: 2001-10-31  
; PRIOR APPLICATION NUMBER: 60/244,524  
; PRIOR FILING DATE: 2000-10-31  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 5  
; LENGTH: 3407  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-002-593-5

Query Match	100.0%	Score 2184;	DB 3;	Length 3407;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 2184;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AGTCGCAAACTTACACTCTTAACCTGATTCTTAAAAAATCTTATAGACTGAAGTTATAC	60	
Db	190	AGTCGCAAACTTACACTCTTAACCTGATTCTTAAAAAATCTTATAGACTGAAGTTATAC	249	
Qy	61	TCCTTAAAGATGGTTTCAGATCATGAATATCTCTACAAACAAGAAATAATATCTTGGTA	120	
Db	250	TCCTTAAAGATGGTTTCAGATCATGAATATCTCTACAAACAAGAAATAATATCTTGGTA	309	
Qy	121	TTCAATGCTGAATATGGAACAGCTCAGTTTTCTTGGAGAACAGTACATCTTGGAGTTTT	180	
Db	310	TTCAATGCTGAATATGGAACAGCTCAGTTTTCTTGGAGAACAGTACATCTTGGAGTTTT	369	
Qy	181	GGACATTTCTATCAATGATTAATTCATATCTCTGATGGGCAAGTTATCTCTTTAGAAATAC	240	
Db	370	GGACATTTCTATCAATGATTAATTCATATCTCTGATGGGCAAGTTATCTCTTTAGAAATAC	429	
Qy	241	AACCTACCTGAGCAATGGGCAATCTCTACACAGCTTCATATGACATTTATGATTTAAAT	300	
Db	430	AACCTACCTGAGCAATGGGCAATCTCTACACAGCTTCATATGACATTTATGATTTAAAT	489	
Qy	301	AAAAGGCGAGCTGATTACAGAAGAGAGGATTCCAAACAACACACAGTGGGTACATGGTCA	360	





QY 2041 ATCTCCAAAGCCCTGGTCGATGTTGGAGTGGATTTCCAGGCAATGTTGGTATATCATGAA 2100  
Db 2230 ATCTCCAAAGCCCTGGTCGATGTTGGAGTGGATTTCCAGGCAATGTTGGTATATCATGAA 2289  
QY 2101 GACCATTGGAATAGCTAGCAGCAGACACACCAACATATATATACCCACATGAGCCACTTC 2160  
Db 2290 GACCATTGGAATAGCTAGCAGCAGACACCAACATATATATATACCCACATGAGCCACTTC 2349  
QY 2161 ATAAACAATGTTCTCTTTACT 2184  
Db 2350 ATAAACAATGTTCTCTTTACT 2373  
RESULT 3  
US-10-423-714-5  
; Sequence 5, Application US/10423714  
; Patent No. 6887679  
; GENERAL INFORMATION:  
; APPLICANT: Vanderbilt University  
; APPLICANT: Brown, Nancy J.  
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTIN  
; FILE REFERENCE: Acty Docket No. 6887679 1242/48/2  
; CURRENT APPLICATION NUMBER: US/10/423,714  
; CURRENT FILING DATE: 2003-04-25  
; PRIOR APPLICATION NUMBER: 60/244,524  
; PRIOR FILING DATE: 2000-10-31  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 3407  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-423-714-5  
Query Match 100.0%; Score 2184; DB 3; Length 3407;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGTCGCAAACTTTACACTCTAACTGATTAATCTTAAAAAATCTTATAGACTGAAGTTATAC 60  
Db 190 AGTCGCAAACTTTACACTCTAACTGATTAATCTTAAAAAATCTTATAGACTGAAGTTATAC 249  
QY 61 TCTTAAAGATGGAATTTGAGATCATGATATCTCTACAAACAGAAAAATATATCTTGGTA 120  
Db 250 TCTTAAAGATGGAATTTGAGATCATGATATCTCTACAAACAGAAAAATATATCTTGGTA 309  
QY 121 TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 180  
Db 310 TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 369  
QY 181 GGACATTTCTATCAATGATTAATCAATATCTCTGATGGCAGTTTATTTCTTTAGAAATAC 240  
Db 370 GGACATTTCTATCAATGATTAATCAATATCTCTGATGGCAGTTTATTTCTTTAGAAATAC 429  
QY 241 AACTAGTGAAGCAATGGAGGCAATTCCTACAGCTTTCATATGACATTTATGATTTAAAT 300  
Db 430 AACTAGTGAAGCAATGGAGGCAATTCCTACAGCTTTCATATGACATTTATGATTTAAAT 489  
QY 301 AAAAGCAGCTGATTAACAGAGAGGATTTCCAAACACACACAGTGGGTCCATGGTCA 360  
Db 490 AAAAGCAGCTGATTAACAGAGAGGATTTCCAAACACACACAGTGGGTCCATGGTCA 549  
QY 361 CCAGTGGGTCAATAAATGGCATATGTTTGGAAACATGACATTTATGTTAAAAATGAAACCA 420  
Db 550 CCAGTGGGTCAATAAATGGCATATGTTTGGAAACATGACATTTATGTTAAAAATGAAACCA 609  
QY 421 AATTTACCAAGTTTACAGATTCAGATCGGAGGAGAAAGAAATATATATATATGGAATA 480  
Db 610 AATTTACCAAGTTTACAGATTCAGATCGGAGGAGAAAGAAATATATATATATGGAATA 669  
QY 481 ACTGACTGGGTTTATGAAGAGGAGTCTTCAGTGCCTACTCTGCTCTGGTGGTCTCCA 540

Db 670 ACTGACTGGGTTTATGAAGAGGAAGTCTTCAGTGCCTACTCTGCTCTGTGGTGGTCTCCA 729  
QY 541 AACGCGCACTTTTATAGCATATATGCCCAATTTAACGACACAGAAAGTCCCACTTATTTGAATAC 600  
Db 730 AACGCGCACTTTTATAGCATATATGCCCAATTTAACGACACAGAAAGTCCCACTTATTTGAATAC 789  
QY 601 TCCTTCTACTCTGATGAGTCACTGCAGTACCCAAAGACTGTACGGGTTCCATATCCAAAG 660  
Db 790 TCCTTCTACTCTGATGAGTCACTGCAGTACCCAAAGACTGTACGGGTTCCATATCCAAAG 849  
QY 661 GCAGGAGCTGTGAATCCAACTGTAAAGTCTTTTGTGTAATAACAGACTCTCTCAGCTCA 720  
Db 850 GCAGGAGCTGTGAATCCAACTGTAAAGTCTTTTGTGTAATAACAGACTCTCTCAGCTCA 909  
QY 721 GTCACCAATGCAATTTCCATACAAATCACTGCTCTGCTCTCTATGTTGATAGGGATCAC 780  
Db 910 GTCACCAATGCAATTTCCATACAAATCACTGCTCTGCTCTCTATGTTGATAGGGATCAC 969  
QY 781 TACTTGTGATGTGACATGGGCAACACAGAAAGAAATTTCTTCAGTGGCTCAGGAGG 840  
Db 970 TACTTGTGATGTGACATGGGCAACACAGAAAGAAATTTCTTCAGTGGCTCAGGAGG 1029  
QY 841 ATTCAAGAACTATTTCGGTCAATGATATTTGTGATGATGAATCCAGTGGAAAGATGGAAC 900  
Db 1030 ATTCAAGAACTATTTCGGTCAATGATATTTGTGATGATGAATCCAGTGGAAAGATGGAAC 1089  
QY 901 TGCTTAGTGGCAACGCAACACATTTGAAATAGTACTACTGGCTGGTGGAAAGATTTAGG 960  
Db 1090 TGCTTAGTGGCAACGCAACACATTTGAAATAGTACTACTGGCTGGTGGAAAGATTTAGG 1149  
QY 961 CTTTCAGAACCTCATTTTACCCTTGATGATTAATAGCTCTCAAGATCATCAGCAATGAA 1020  
Db 1150 CTTTCAGAACCTCATTTTACCCTTGATGATTAATAGCTCTCAAGATCATCAGCAATGAA 1209  
QY 1021 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATAAAAAAGACTGCAATTATTTACA 1080  
Db 1210 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATAAAAAAGACTGCAATTATTTACA 1269  
QY 1081 AAAGGCACTGGGAAAGTCAATGGGATAGAAGCTCTAACAGTGATATCTACTACTACT 1140  
Db 1270 AAAGGCACTGGGAAAGTCAATGGGATAGAAGCTCTAACAGTGATATCTACTACTACT 1329  
QY 1141 AGTAATGAATATAAAGGAATGCCAGGAGGAAGAAATCTTTATAAAATCCAACTTATTGAC 1200  
Db 1330 AGTAATGAATATAAAGGAATGCCAGGAGGAAGAAATCTTTATAAAATCCAACTTATTGAC 1389  
QY 1201 TATACAAAGTGACATGCTCAGTTTGTGAGCTGAAATCCGGAAGGTTGTCAGTACTATTCT 1260  
Db 1390 TATACAAAGTGACATGCTCAGTTTGTGAGCTGAAATCCGGAAGGTTGTCAGTACTATTCT 1449  
QY 1261 GTGTCATTTCAGTAAGAGCGGAAGTATTATCAGCTGAGATGTTCCGGTCTGCTGCTGCC 1320  
Db 1450 GTGTCATTTCAGTAAGAGCGGAAGTATTATCAGCTGAGATGTTCCGGTCTGCTGCTGCC 1509  
QY 1321 CTCTACTCTACACAGCAGCGTGAATGATAAAGGCTGAGAGTCTCTGGAAGCAATTTCA 1380  
Db 1510 CTCTACTCTACACAGCAGCGTGAATGATAAAGGCTGAGAGTCTCTGGAAGCAATTTCA 1569  
QY 1381 GCTTTGGGATAAAATGCTCCAGAAATGTCAGATGCTCCCAAAAAAAGTGGACTTTCAATT 1440  
Db 1570 GCTTTGGGATAAAATGCTCCAGAAATGTCAGATGCTCCCAAAAAAAGTGGACTTTCAATT 1629  
QY 1441 TTGAATGAAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCTATTTGATATAATCCAAG 1500  
Db 1630 TTGAATGAAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCTATTTGATATAATCCAAG 1689  
QY 1501 AAATATCTCTACTATTAAGTGTGATGAGGCCCCATGAGTCAAAAGACAGACACTGTC 1560  
Db 1690 AAATATCTCTACTATTAAGTGTGATGAGGCCCCATGAGTCAAAAGACAGACACTGTC 1749  
QY 1561 TTCAGACTGAACCTGGGCCACTTTACCTTGAAGCACAGAAAAATTTATAGTAGCTAGCTTTT 1620

Db	1750	TTCAGACTGAAC	CTGGGCCACTTACCTTGC	AAGCAGAAAA	CAATTATAGTAGCTAGCTTT	1809
Qy	1621	GATGCCAGAGGA	AGTGGTTTACCAAGCAGATA	AGATCATGCATGC	CAATCAACAGAACTG	1680
Db	1810	GATGCCAGAGGA	AGTGGTTTACCAAGCAGATA	AGATCATGCATGC	CAATCAACAGAACTG	1869
Qy	1681	GGAACTTTGAAG	TTGAAGATCAAAATTGAAG	CAGCCAGACAATTTTCA	AAAAATGGGATTT	1740
Db	1870	GGAACTTTGAAG	TTGAAGATCAAAATTGAAG	CAGCCAGACAATTTTCA	AAAAATGGGATTT	1929
Qy	1741	GTGGACAACAAC	GAAATGTCAAATTTTGGG	CGCTGTCATATGGAG	GGGTACGTAACTCAATG	1800
Db	1930	GTGGACAACAAC	GAAATGTCAAATTTTGGG	CGCTGTCATATGGAG	GGGTACGTAACTCAATG	1989
Qy	1801	GTCTCTGGGAT	CGGGAAGTGGCGTGTTC	CAAGTGTGGAATAG	CCGTCGGCGCTGTATCCC	1860
Db	1990	GTCTCTGGGAT	CGGGAAGTGGCGTGTTC	CAAGTGTGGAATAG	CCGTCGGCGCTGTATCCC	2049
Qy	1861	TGGGAGTACTAT	GACTCAGTGTACACAGAA	CGTTTACATGGGTC	TCCCAACTCCAGAAGAC	1920
Db	2050	TGGGAGTACTAT	GACTCAGTGTACACAGAA	CGTTTACATGGGTC	TCCCAACTCCAGAAGAC	2109
Qy	1921	AACCTTGACCAT	TACAGAAATTCACACAGT	CAATGAGCAGAGCTG	MAAAATTTTAAACAAGTT	1980
Db	2110	AACCTTGACCAT	TACAGAAATTCACACAGT	CAATGAGCAGAGCTG	MAAAATTTTAAACAAGTT	2169
Qy	1981	GAGTACCTCTCT	TATTCATGGAAACAGCAG	ATGATAACGTTTCACTTT	CAGCAGTCAGCTCAG	2040
Db	2170	GAGTACCTCTCT	TATTCATGGAAACAGCAG	ATGATAACGTTTCACTTT	CAGCAGTCAGCTCAG	2229
Qy	2041	ATCTCCAAAGCC	CTGGTGCATGTGGAGTGG	ATTTCCAGGCCAATGTGGT	TATCTAGTGA	2100
Db	2230	ATCTCCAAAGCC	CTGGTGCATGTGGAGTGG	ATTTCCAGGCCAATGTGGT	TATCTAGTGA	2289
Qy	2101	GACCATGGAAT	AGCTAGCAGACACACCA	CAATATATATACCC	ACATGAGCCACTTC	2160
Db	2290	GACCATGGAAT	AGCTAGCAGACACCA	CAATATATATACCC	ACATGAGCCACTTC	2349
Qy	2161	ATAAAACAATG	TTTCTCTTTTACCT			2184
Db	2350	ATAAAACAATG	TTTCTCTTTTACCT			2373

RESULT 4

US-09-949-016-4579

Sequence 4579, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 4579

LENGTH: 3419

TYPE: DNA

ORGANISM: Human

US-09-949-016-4579

Query Match 99.9%; Score 2182.4; DB 3; Length 3419;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 218; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 AGTGGCAAACTTCACTCTAAGTATTACTTAAAAAATACTTATAGACTGAAGTTATAC 60

191	DB	AGTCGCAAAACTTACACTCTAACTGATTA	CTCTAAAAAATACTTATAGACTGAAGTTATAC	250
61	QY	TCCTTTAAGATGGATTTCCAGATCATGAATATCTCTACAAACAAGAAATAATAATCTTCGTGA	120	
251	DB	TCCTTTAAGATGGATTTCCAGATCATGAATATCTCTACAAACAAGAAATAATAATCTTCGTGA	310	
121	QY	TTCAATGCTGAATATGGAAACAGCTCAGTCTTTTCTTGAGAACAGTACATATTTGATGAGTTT	180	
311	DB	TTCAATGCTGAATATGGAAACAGCTCAGTCTTTTCTTGAGAACAGTACATATTTGATGAGTTT	370	
181	QY	GGACATTTCTATCAATGATTAATTTCAATATCTCTGATGGGCAAGTTTATTTCTTTAGAAATAC	240	
371	DB	GGACATTTCTATCAATGATTAATTTCAATATCTCTGATGGGCAAGTTTATTTCTTTAGAAATAC	430	
241	QY	AATCTACGTGAAGCAATCGAGGCATTTCTCTACACAGCTTTCATATGACATTTTATGATTTAAAT	300	
431	DB	AATCTACGTGAAGCAATCGAGGCATTTCTCTACACAGCTTTCATATGACATTTTATGATTTAAAT	490	
301	QY	AAAAAGGCAGCTGATTAACAGAGAGAGGATTTCCAAACAACACACAGTGGGTCACATGTGTCA	360	
491	DB	AAAAAGGCAGCTGATTAACAGAGAGAGGATTTCCAAACAACACACAGTGGGTCACATGTGTCA	550	
361	QY	CCAGTGGGTCAATAAATTTGGCATATGTTTGGAACTATTTATGTTAAAAATTGAACCA	420	
551	DB	CCAGTGGGTCAATAAATTTGGCATATGTTTGGAACTATTTATGTTAAAAATTGAACCA	610	
421	QY	AAATTTTCAACAGTTTACAGAAATCACATGGACGGGGAAAGAAAGATATAATATATATGGAATA	480	
611	DB	AAATTTTCAACAGTTTACAGAAATCACATGGACGGGGAAAGAAAGATATAATATATATGGAATA	670	
481	QY	ACTGACTGGGTTTATGAAGAGGAGTCTTCAGTGCCTTACTCTGCTCTGTGGTGGTCTCCA	540	
671	DB	ACTGACTGGGTTTATGAAGAGGAGTCTTCAGTGCCTTACTCTGCTCTGTGGTGGTCTCCA	730	
541	QY	AAACGGCACTTTTTTAGCATATGCCAATTTAAACGACAACAGAGTCCCACTTATTTGAATATAC	600	
731	DB	AAACGGCACTTTTTTAGCATATGCCAATTTAAACGACAACAGAGTCCCACTTATTTGAATATAC	790	
601	QY	TCTTCTTACTCTGATGAGTCACTGCAGTACCCAAAGACTGTACGGGTTCATATCCCAAG	660	
791	DB	TCTTCTTACTCTGATGAGTCACTGCAGTACCCAAAGACTGTACGGGTTCATATCCCAAG	850	
661	QY	GCAGGAGCTGTGAATCCAACTGTAAAGTCTTTTGTGTTAAATACAGACTCTCTCAGCTCA	720	
851	DB	GCAGGAGCTGTGAATCCAACTGTAAAGTCTTTTGTGTTAAATACAGACTCTCTCAGCTCA	910	
721	QY	GTCAACCAATGCAACTTCCATACAAATCACTGCTCTCTGTTCTATGTTGATAGGGGATCAC	780	
911	DB	GTCAACCAATGCAACTTCCATACAAATCACTGCTCTCTGTTCTATGTTGATAGGGGATCAC	970	
781	QY	TACTTGTGTGATGTGATCGGGCAACACAGAAGAATTTCTTTGTCAGTGGCTCAGGAGG	840	
971	DB	TACTTGTGTGATGTGATCGGGCAACACAGAAGAATTTCTTTGTCAGTGGCTCAGGAGG	1030	
841	QY	ATTTCAGAACTATTTCGGTTCATGGATATTGTTGACTATGATGAATCCAGTGGAAAGTGAAC	900	
1031	DB	ATTTCAGAACTATTTCGGTTCATGGATATTGTTGACTATGATGAATCCAGTGGAAAGTGAAC	1090	
901	QY	TGCTTTATGTGCAACGGCAACACATTTGAAAATGAGTACTACTGGCTGGGTTCGGAAGATTTAGG	960	
1091	DB	TGCTTTATGTGCAACGGCAACACATTTGAAAATGAGTACTACTGGCTGGGTTCGGAAGATTTAGG	1150	
961	QY	CCTTTCAGAACTCATTTTACCTTTGATGGTAAATAGCTTCTTACAAAGATCATCAGCAATGAA	1020	
1151	DB	CCTTTCAGAACTCATTTTACCTTTGATGGTAAATAGCTTCTTACAAAGATCATCAGCAATGAA	1210	
1021	QY	GAAGGTTACAGACACATTTTGTCTATTTTCCAAATAGATAAAAAGACTGCACATTTTATTACA	1080	
1211	DB	GAAGGTTACAGACACATTTTGTCTATTTTCCAAATAGATAAAAAGACTGCACATTTTATTACA	1270	
1081	QY	AAAGGCACTGGGAAGTCACTCGGGATAGAGCTCTTAACCAAGTGAATATCTTATCTACTACATT	1140	



Db 1271 AAAGGCACCTGGAGTCACTCGGATAGAAGCTCTAAACAGTGATTATCTATCTACTACATT 1330  
 Qy 1141 AGTAATGAATATAAGGAATCCAGGAGGAAGTCTTTATATAAATCCAACTTATTGAC 1200  
 Db 1331 AGTAATGAATATAAGGAATCCAGGAGGAAGTCTTTATATAAATCCAACTTATTGAC 1390  
 Qy 1201 TATACAAAAGTGCATGCTCAGTTGTGAGCTGAATCCGAAAGGTGTCAGTACTATTCT 1260  
 Db 1391 TATACAAAAGTGCATGCTCAGTTGTGAGCTGAATCCGAAAGGTGTCAGTACTATTCT 1450  
 Qy 1261 GTGTCACTCAGTAAAGAGCGAAGTATTATCAGCTGAGATGTTCCGGTCTCTGTCGCC 1320  
 Db 1451 GTGTCACTCAGTAAAGAGCGAAGTATTATCAGCTGAGATGTTCCGGTCTCTGTCGCC 1510  
 Qy 1321 CTCCTATCTACACAGCAGCGTGAATGATTAAGGCTGAGAGTCTCTGGAAGCAATTC 1380  
 Db 1511 CTCCTATCTACACAGCAGCGTGAATGATTAAGGCTGAGAGTCTCTGGAAGCAATTC 1570  
 Qy 1381 GCTTTGGATATAATGCTGCAGATGTCAGATGCCCTCCAAAAAACTGGACTTCATTATT 1440  
 Db 1571 GCTTTGGATATAATGCTGCAGATGTCAGATGCCCTCCAAAAAACTGGACTTCATTATT 1630  
 Qy 1441 TTGAATGAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGTATAATCCAG 1500  
 Db 1631 TTGAATGAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGTATAATCCAG 1690  
 Qy 1501 AAATATCCTCTACTATTAGATGTATGAGGCGCATGTAGTCAAAAAGCAGACACTGTC 1560  
 Db 1691 AAATATCCTCTACTATTAGATGTATGAGGCGCATGTAGTCAAAAAGCAGACACTGTC 1750  
 Qy 1561 TTCAAGCTCACTGGGCTACCTTACCTTGAAGCAGACAGAAAATATATAGTACTAGCTTT 1620  
 Db 1751 TTCAAGCTCACTGGGCTACCTTACCTTGAAGCAGACAGAAAATATATAGTACTAGCTTT 1810  
 Qy 1621 GATGCGAGAGGAGTGGTTTACCAAGGAGATAAGATCATGCTGCAATCAACAGAGACTG 1680  
 Db 1811 GATGCGAGAGGAGTGGTTTACCAAGGAGATAAGATCATGCTGCAATCAACAGAGACTG 1870  
 Qy 1681 GGAACATTTGAAGTTGAAGATCAATTTGAAGCAGCAGACAATTTTCAAAAATGGGATTT 1740  
 Db 1871 GGAACATTTGAAGTTGAAGATCAATTTGAAGCAGCAGACAATTTTCAAAAATGGGATTT 1930  
 Qy 1741 GTGACACAAACGAATTCGAATTTGGGCTGTGTATGAGGAGTACGTAACTCAATG 1800  
 Db 1931 GTGACACAAACGAATTCGAATTTGGGCTGTGTATGAGGAGTACGTAACTCAATG 1990  
 Qy 1801 GTCTGGGATCGGGAAGTGGCGTTCAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG 1860  
 Db 1991 GTCTGGGATCGGGAAGTGGCGTTCAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG 2050  
 Qy 1861 TGGGAGTACTATGACTCAGTGATACAGAAAGTTATAGGGTCTCCAACTCCAGAGAC 1920  
 Db 2051 TGGGAGTACTATGACTCAGTGATACAGAAAGTTATAGGGTCTCCAACTCCAGAGAC 2110  
 Qy 1921 AACCTTGACCATTACAGAAATTCACAGTCATGAGCAGAGCTGAAAATTTTAAACAGTT 1980  
 Db 2111 AACCTTGACCATTACAGAAATTCACAGTCATGAGCAGAGCTGAAAATTTTAAACAGTT 2170  
 Qy 1981 GAGTACTCTCTTATTATGGAACAGCAGATGATTAAGTTTCACTTTTCCAGCAGTCAAG 2040  
 Db 2171 GAGTACTCTCTTATTATGGAACAGCAGATGATTAAGTTTCACTTTTCCAGCAGTCAAG 2230  
 Qy 2041 ATCTCAAAAGCCCTGGTGTGATGTTGAGTGGATTTCCAGGCAATGTGGTATATGATGAA 2100  
 Db 2231 ATCTCAAAAGCCCTGGTGTGATGTTGAGTGGATTTCCAGGCAATGTGGTATATGATGAA 2290  
 Qy 2101 GACCATGGAATAGTACGACGACAGCAGACCAACATATATATATATATATATATATATAT 2160  
 Db 2291 GACCATGGAATAGTACGACGACAGCAGACCAACATATATATATATATATATATATATAT 2350  
 Qy 2161 ATAAAAAATGTTTCTCTTTACCT 2184  
 Db 2351 ATAAAAAATGTTTCTCTTTACCT 2374

RESULT 5  
 US-09-023-655-1026  
 ; Sequence 1026, Application US/09023655  
 ; Patent No. 6607879  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cocks, Benjamin G.  
 ; APPLICANT: Susan G. Stuart  
 ; APPLICANT: Jeffrey J. Seilhamer  
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE  
 ; TITLE OF INVENTION: EXPRESSION  
 ; NUMBER OF SEQUENCES: 1508  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
 ; STREET: 3174 PORTER DRIVE  
 ; CITY: PALO ALTO  
 ; STATE: CALIFORNIA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/023,655  
 ; FILING DATE: HEREWITH  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Zeller, Karen J.  
 ; REGISTRATION NUMBER: 37,071  
 ; REFERENCE/DOCKET NUMBER: PA-0001 US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (650) 855-0555  
 ; TELEFAX: (650) 845-4166  
 ; INFORMATION FOR SEQ ID NO: 1026:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2924 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: GENBANK  
 ; CLONE: g180082  
 ; US-09-023-655-1026

Query Match 99.9%; Score 2180.8; DB 3; Length 2924;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 2182; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AGTCCCAAACTTACACTTAACCTGATTACTTAAAAATACTTATAGACTGAAGTTATAC 60  
 Db 125 AGTCCCAAACTTACACTTAACCTGATTACTTAAAAATACTTATAGACTGAAGTTATAC 184  
 Qy 61 TCCTTAAAGATGGATTTCAGATCATGAATATCTCTACAAACAGAAAAATAATATCTTGGTA 120  
 Db 185 TCCTTAAAGATGGATTTCAGATCATGAATATCTCTACAAACAGAAAAATAATATCTTGGTA 244  
 Qy 121 TTCATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 180  
 Db 245 TTCATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 304  
 Qy 181 GGACATTTCTCAATGATTATTCATATCTCTCTGATGGCAGTTTATCTCTTAGAATAC 240  
 Db 305 GGACATTTCTCAATGATTATTCATATCTCTCTGATGGCAGTTTATCTCTTAGAATAC 364  
 Qy 241 AACTACGTGAAGCAATGGAGGCAATTCCTACACAGCTTCATATGACATTTATGATTTAAAT 300

365	Db	AACTACGTGAAGCAATGGAGGCATTCCTACACAGCTTCATATGACATTTATGATTTAAAT	424
301	Qy	AAAAAGCAGCTGATTTACAGAAAGAGAGGANTTCAAACAAACACACAGTGGGTCCACATGGTCA	360
425	Db	AAAAGGCAGCTGATTTACAGAAAGAGAGGANTTCAAACAAACACACAGTGGGTCCACATGGTCA	484
361	Qy	CCAGTGGGTCTAATAATTTGGGCATATGTTTGGCAACAATGACATTTATGTTAAAATTTGAACCA	420
485	Db	CCAGTGGGTCTAATAATTTGGGCATATGTTTGGCAACAATGACATTTATGTTAAAATTTGAACCA	544
421	Qy	AAATTTACCAGGTACAGAAATCACATGGAACGGGGAAGAAGATATAAATATATAAATGGAATA	480
545	Db	AAATTTACCAGGTACAGAAATCACATGGAACGGGGAAGAAGATATAAATATATAAATGGAATA	604
481	Qy	ACTGACTGGGTTTATGAAGAGAGAACTCTTCAGTGCCTACTCTGCTGTGTGGTGTCTCCCA	540
605	Db	ACTGACTGGGTTTATGAAGAGAGAACTCTTCAGTGCCTACTCTGCTGTGTGGTGTCTCCCA	664
541	Qy	AACGGCACTTTTATGACATATGCCCAATTTAAACGACACAGAAAGTCCCACTTATTTGAATAC	600
665	Db	AACGGCACTTTTATGACATATGCCCAATTTAAACGACACAGAAAGTCCCACTTATTTGAATAC	724
601	Qy	TCCTTCTACTCTGATGATGACTCGACGTACCCAAAGACTGTACGGGTTCATATCCAAAG	660
725	Db	TCCTTCTACTCTGATGATGACTCGACGTACCCAAAGACTGTACGGGTTCATATCCAAAG	784
661	Qy	GCAGGAGCTGTGAATCCAACTGTAAAGTTCTTTGTGTAAATACAGACTCTCTCAGCTCA	720
785	Db	GCAGGAGCTGTGAATCCAACTGTAAAGTTCTTTGTGTAAATACAGACTCTCTCAGCTCA	844
721	Qy	GTCAACCAATGCAATTCCTCATACAAATCACTGCTCCTCTTCTATGTTGATAGGGGATCAC	780
845	Db	GTCAACCAATGCAATTCCTCATACAAATCACTGCTCCTCTTCTATGTTGATAGGGGATCAC	904
781	Qy	TACTTGTGTGATGTGACATGCGGCAACACAGAAAGAAATTTCTTTGTCAGTGGCTCAGAGG	840
905	Db	TACTTGTGTGATGTGACATGCGGCAACACAGAAAGAAATTTCTTTGTCAGTGGCTCAGAGG	964
841	Qy	ATTCAGAACTATTCCGGTCAATGGAATATTTGTGACTATGATGAATCCAGTGGAAAGATGGAAC	900
965	Db	ATTCAGAACTATTCCGGTCAATGGAATATTTGTGACTATGATGAATCCAGTGGAAAGATGGAAC	1024
901	Qy	TGCTTAGTGGCACGGCAACACATTTGAAATGAGTACTACTGCTGGGTGGTGGAAAGATTTAGG	960
1025	Db	TGCTTAGTGGCACGGCAACACATTTGAAATGAGTACTACTGCTGGGTGGTGGAAAGATTTAGG	1084
961	Qy	CCCTTCAGAACTCAATTTTACCCCTTGATGTTAATAGTCTTCTACAAGATCATCAGCAATGAA	1020
1085	Db	CCCTTCAGAACTCAATTTTACCCCTTGATGTTAATAGTCTTCTACAAGATCATCAGCAATGAA	1144
1021	Qy	GAAGGTTACAGACACATTTGCTATTTCCAAATAGATATAAAAGACTGCACATTTATTATCA	1080
1145	Db	GAAGGTTACAGACACATTTGCTATTTCCAAATAGATATAAAAGACTGCACATTTATTATCA	1204
1081	Qy	AAAGGCACCTGGGAAGTCATCGGGATAGAAGCTCTTAACCAAGTATTTACTATCTACATT	1140
1205	Db	AAAGGCACCTGGGAAGTCATCGGGATAGAAGCTCTTAACCAAGTATTTACTATCTACATT	1264
1141	Qy	AGTAAATGAATATAAAGGAATGCCAGGAGGAAGGAATCTTTTATAAAATCCAACTTATTGAC	1200
1265	Db	AGTAAATGAATATAAAGGAATGCCAGGAGGAAGGAATCTTTTATAAAATCCAACTTATTGAC	1324
1201	Qy	TATACAAAAGTGACATCGCTCAGTTGTGAGCTGAATCCGGAAAGGTGTCAAGTACTATTCT	1260
1325	Db	TATACAAAAGTGACATCGCTCAGTTGTGAGCTGAATCCGGAAAGGTGTCAAGTACTATTCT	1384
1261	Qy	GTGTCTATTAGTAAAGGGCGAAGTATATATCAGCTGAGATGTTCCGGTCTCGGTCTGCCCC	1320
1385	Db	GTGTCTATTAGTAAAGGGCGAAGTATATATCAGCTGAGATGTTCCGGTCTCGGTCTGCCCC	1444
1321	Qy	CTCTATCTCTTACACAGAGCGGTGAATGATAAAGGGCTGAGAGTCTCGGAAGACAATTTCA	1380
1445	Db	CTCTATCTCTTACACAGAGCGGTGAATGATAAAGGGCTGAGAGTCTCGGAAGACAATTTCA	1504

```

RESULT 6
PCT-US93-07923-1
; Sequence 1, Application PC/TUS9307923
; GENERAL INFORMATION:
;
; APPLICANT: Morimoto, Chikao
; APPLICANT: Schlossman, Stuart P.
; APPLICANT: Tanaka, Toshiaki
; TITLE OF INVENTION: HUMAN CD26 AND METHODS FOR USE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.

```

```

; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: WordPerfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07923
; FILING DATE: 19930819
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934,162
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: 07/832,211
; FILING DATE: 08-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2924
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; PCT-US93-07923-1

Query Match 99.9%; Score 2180.8; DB 6; Length 2924;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2182; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AGTCGCAAACTTACACTCTAATGATTAATCTTAAATAAATCTTATAGAGTGAAGTTATAC 60
DB 125 AGTCGCAAACTTACACTCTAATGATTAATCTTAAATAAATCTTATAGAGTGAAGTTATAC 184
QY 61 TCCTTAAGATGGATTTAGATCATGATATCTCTACAAACAAAGAAATATATCTTGGTA 120
DB 185 TCCTTAAGATGGATTTAGATCATGATATCTCTACAAACAAAGAAATATATCTTGGTA 244
QY 121 TTCATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 180
DB 245 TTCATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 304
QY 181 GGACATTTCTATCAATGATTTCAATATCTCTGATGGGAGTTTATTTCTTTAGAAATAC 240
DB 305 GGACATTTCTATCAATGATTTCAATATCTCTGATGGGAGTTTATTTCTTTAGAAATAC 364
QY 241 AACTAGTGAAGCAATGGAGGCAATTCCTTACAGCTTTCATATGACATTTATGATTTAAAT 300
DB 365 AACTAGTGAAGCAATGGAGGCAATTCCTTACAGCTTTCATATGACATTTATGATTTAAAT 424
QY 301 AAAAGCAGCTGATTTACAGAGAGGATTCCTTCAACAAACACAGCTGGTTCATGGTCA 360
DB 425 AAAAGCAGCTGATTTACAGAGAGGATTCCTTCAACAAACACAGCTGGTTCATGGTCA 484
QY 361 CCAAGTGGGTTCATAAAATGGCATATGTTTGGAAACAAATGACATTTATGTTAAATTTGAACCA 420
DB 485 CCAAGTGGGTTCATAAAATGGCATATGTTTGGAAACAAATGACATTTATGTTAAATTTGAACCA 544
QY 421 AATTACCAAGTTTACAGATTCATGAGCGGGGAAAGAGATATATATATATATGGAATA 480
DB 545 AATTACCAAGTTTACAGATTCATGAGCGGGGAAAGAGATATATATATATATGGAATA 604
QY 481 ACTGACTGGTTTATCAAGGAGGAGTCTTTCAGTGGCTACTCTGCTGTGGTGGTCTCCA 540
DB 605 ACTGACTGGTTTATCAAGGAGGAGTCTTTCAGTGGCTACTCTGCTGTGGTGGTCTCCA 664
QY 541 AACGGCACTTTTATAGCATATGCCCAATTTTAAACGACACAGAAAGTCCCACTTATGAAATAC 600

```

```

DB 665 AACGGCACTTTTATAGCATATGCCCAATTTTAAACGACACAGAAAGTCCCACTTATTGAATAC 724
QY 601 TCCTTCTACTCTGATGAGTCACTGAGTACCCAAAGACTGTACGGGTTCCCATATCCAAAG 660
DB 725 TCCTTCTACTCTGATGAGTCACTGAGTACCCAAAGACTGTACGGGTTCCCATATCCAAAG 784
QY 661 GCAGGAGCTGTGAATCCAACTGTAAGTCTTTTGTGTAATAACAGACTCTCTCAGCTCA 720
DB 785 GCAGGAGCTGTGAATCCAACTGTAAGTCTTTTGTGTAATAACAGACTCTCTCAGCTCA 844
QY 721 GTCCAAATGCAACTTCCATACAAATCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
DB 845 GTCCAAATGCAACTTCCATACAAATCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 904
QY 781 TACTTGTGTGATGTGACATGGGCAACACAAAGAAAGAAATTTCTTTTCAGCTGAGGAGG 840
DB 905 TACTTGTGTGATGTGACATGGGCAACACAAAGAAAGAAATTTCTTTTCAGCTGAGGAGG 964
QY 841 ATTCCAGAACTATTTCGGTTCATGGATATTGTGACTATGATCCAGTCCAGTCCAGTCCAGT 900
DB 965 ATTCCAGAACTATTTCGGTTCATGGATATTGTGACTATGATCCAGTCCAGTCCAGTCCAGT 1024
QY 901 TGTCTAGTGGCACGGCAACACATTCGAAATGAGTACTTCTGCTGGTGGTGGTGGTGGTGGT 960
DB 1025 TGTCTAGTGGCACGGCAACACATTCGAAATGAGTACTTCTGCTGGTGGTGGTGGTGGTGGT 1084
QY 961 CTTTCAGAACTCTCATTTTACCTTTCATGGTAAATAGCTTCTCAAGTCAATCAGCAATGAA 1020
DB 1085 CTTTCAGAACTCTCATTTTACCTTTCATGGTAAATAGCTTCTCAAGTCAATCAGCAATGAA 1144
QY 1021 GAAGGTTACAGACACATTTGCTTATTTCCAAATAGATAAAAGACTGCAATTTTATTTACA 1080
DB 1145 GAAGGTTACAGACACATTTGCTTATTTCCAAATAGATAAAAGACTGCAATTTTATTTACA 1204
QY 1081 AAAGCACCCTGGGAGTCACTCGGATAGAAGCTCTAACAGTGAATTAATCTATCTATCTAT 1140
DB 1205 AAAGCACCCTGGGAGTCACTCGGATAGAAGCTCTAACAGTGAATTAATCTATCTATCTAT 1264
QY 1141 AGTAATGAAATATAAGGAATGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
DB 1265 AGTAATGAAATATAAGGAATGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1324
QY 1201 TATACAAAGTGAACATGCTCAGTTGTGAGCTGAAATCCGGAAGGAGTGTCCAGTACTATTCT 1260
DB 1325 TATACAAAGTGAACATGCTCAGTTGTGAGCTGAAATCCGGAAGGAGTGTCCAGTACTATTCT 1384
QY 1261 GTGTCATTTCACTGTAAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320
DB 1385 GTGTCATTTCACTGTAAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1444
QY 1321 CTCTATCTCTACACAGCAGCGTGAAATGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380
DB 1445 CTCTATCTCTACACAGCAGCGTGAAATGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1504
QY 1381 GCTTTGGGATATAATGCTGCAAGATGCTCCAGATGCTCCAGATGCTCCAGATGCTCCAGAT 1440
DB 1505 GCTTTGGGATATAATGCTGCAAGATGCTCCAGATGCTCCAGATGCTCCAGATGCTCCAGAT 1564
QY 1441 TTGAAATGAAACAAAATTTTGGTATCAGATGATTTTGGTATCAGATGATTTTGGTATGAAAT 1500
DB 1565 TTGAAATGAAACAAAATTTTGGTATCAGATGATTTTGGTATCAGATGATTTTGGTATGAAAT 1624
QY 1501 AAATATCTCTCTACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
DB 1625 AAATATCTCTCTACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1684
QY 1561 TTCCAGCTGAACTGGGCGCACTTACCTTCCAGACACAGAAACAAATTTATAGTACTAGCTTT 1620
DB 1685 TTCCAGCTGAACTGGGCGCACTTACCTTCCAGACACAGAAACAAATTTATAGTACTAGCTTT 1744
QY 1621 GATGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680
DB 1745 GATGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1804

```

```
QY 1691 GGACATTGAGTTGAAGATCAAAATTGAAGCAGCCAGACAAATTTTCAAAAATGGGATTT 1740
DB 1805 GGAACATTGAGTTGAAGATCAAAATTGAAGCAGCCAGACAAATTTTCAAAAATGGGATTT 1864
QY 1741 GTGGACAAACAAAGAAATGCAATTTTGGGGCTGGTCATATGAGGGGTACGTAACTCAATG 1800
DB 1865 GTGGACAAACAAAGAAATGCAATTTTGGGGCTGGTCATATGAGGGGTACGTAACTCAATG 1924
QY 1801 GTCCTGGGATCGGGAAGTGGCGTGTTCAGAGTGTGGAATAGCCGTGGCGCCTGTATCCCGG 1860
DB 1925 GTCCTGGGATCAGGAAGTGGCGTGTTCAGAGTGTGGAATAGCCGTGGCGCCTGTATCCCGG 1984
QY 1861 TGGGAGTACTATGACTCAGTGTACACAGAAAGTTATCATGGGTCTCCCAACTCCAGAGAC 1920
DB 1985 TGGGAGTACTATGACTCAGTGTACACAGAAAGTTATCATGGGTCTCCCAACTCCAGAGAC 2044
QY 1921 AACCTTGACCAATACAGAAATTCACAGTCATGAGCAGAGCTGAAATTTTAAACAAAGTT 1980
DB 2045 AACCTTGACCAATACAGAAATTCACAGTCATGAGCAGAGCTGAAATTTTAAACAAAGTT 2104
QY 1981 GAGTACCTCTCTTATTCATGGAACAGCAGATGATACGTTTCACTTTTCCAGCAGTCACTCAG 2040
DB 2105 GAGTACCTCTCTTATTCATGGAACAGCAGATGATACGTTTCACTTTTCCAGCAGTCACTCAG 2164
QY 2041 ATCTCCAAAGCCCTGGTGCATGTTGGAGTGGATTTCCAGGCAATGTGGTATATCTGATGAA 2100
DB 2165 ATCTCCAAAGCCCTGGTGCATGTTGGAGTGGATTTCCAGGCAATGTGGTATATCTGATGAA 2224
QY 2101 GACATGGAATAGCTAGCAGACAGCAGACACCAACATATATATATATATATATATATATAT 2160
DB 2225 GACATGGAATAGCTAGCAGACAGCAGACACCAACATATATATATATATATATATATATAT 2284
QY 2161 ATAAACAATGTTCTCTTTTACCT 2184
DB 2285 ATAAACAATGTTCTCTTTTACCT 2308
```

## RESULT 7

US-08-230-491A-1

; Sequence 1, Application US/08230491A

; Patent No. 5587299

; GENERAL INFORMATION:

; APPLICANT: Rettig, Wolfgang J.; Scanlan, Matthew J.;

; APPLICANT: Garin-Chesa, Pilar; Old, Lloyd J.

; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING FOR

; TITLE OF INVENTION: FIBROBLAST ACTIVATION PROTEIN AND USES

; TITLE OF INVENTION: THEREOF

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FELPE &amp; LYNCH

; STREET: 805 THIRD AVENUE

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: USA

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: DISKETTE 3.5 inch 1.2 MB STORAGE

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: PC-DOS

; SOFTWARE: WORDPERFECT - ASC II

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/230,491A

; FILING DATE: 20-APRIL-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Hanson, No. 5587299man D.

; REGISTRATION NUMBER: 30,946

; REFERENCE/DOCKET NUMBER: LUD 330

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 688-9200

; TELEFAX: (212) 838-3884

; INFORMATION FOR SEQ ID NO: 1:

## ; SEQUENCE CHARACTERISTICS:

; LENGTH: 2815 Base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; US-08-230-491A-1

Query Match 34.4%; Score 751; DB 2; Length 2815;

Best Local Similarity 60.9%; Pred. No. 1e-218;

Matches 1320; Conservative 0; Mismatches 825; Indels 24; Gaps 5;

```
QY 15 CACTCTAATGATTAATCTTAAATAATCTTATAGACTGAAGTTATCTCTTAAGATGAT 74
DB 334 CACACTGAAGGATATTTTAAATGGAACATTTTCTTATAAAACATTTTTCCTCAAACTGGAT 393
QY 75 TTCAGATCATCAATATCTCTAACAACAAGAAATAATCTTGGTATTCAATGCTGAATA 134
DB 394 TTCAGGACAGAGATATCTTCATCAATCTGCGAGATAACATATATAGTACTTTATATATGTA 453
QY 135 TGGAAACAGCTCAGTTTTCTTGGAGAACAGTACATTTTGTAGTGTGGACATCTTATCAAA 194
DB 454 AACAGGACAATCATATACCATTTTGGAGTAATAGAACCATGAAAGTGTGAATGCTTCAA- 512
QY 195 TGATTTATCAATATCTCTGATGGGCGAGTTTATTTCTCTTGAATACAACTAGCTGAAGCA 254
DB 513 --ATTACGGCTTATCACCTGATCGGCAATTTGTATATCTAGAAAAGTGAATTATTTCAAAGCT 570
QY 255 ATGAGGAGCATTCCTACACAGCTTCTATGACATTTATGATTTAAATAAAGCGAGCTGAT 314
DB 571 TTGAGAGTACTCTTACACAGCAATATATATCTATGACCTTAGCAATGGAGAAATTTGT 630
QY 315 TACAGAAGAGAGGAGTTCACAAACAAACACACAGTGGGTCACTGGTCCACAGTGGGTCTATAA 374
DB 631 AAGAGGAAATGAGCTTCTCTGCTCAATTCAGTATTTATGCTGGTGGCTGTGGGAGTAA 690
QY 375 ATTGGCATATGTTTGGAAACAATGACATTTATGTTTAAATTTGAACCAAAATTTTACCAAGTTA 434
DB 691 ATTAGCATATGCTCTATCAAAACAATATCTATTTGAAACAAAGACACGAGGAGATCCACCTTT 750
QY 435 CAGAAATCACATGGAAGGGAAGAGATATATATATGAAATTAATGAAATTAATGAAATTTTAA 494
DB 751 TCAAAATCACTTAATGGAAGAAATAAAATATTTAATGAAATTTTAAATGAAATTTTAAATG 810
QY 495 TGAAGAGAGAGTCTTTCAGTGCCTACTCTGCTCTGTGGTGGTCTCCAAACGGCAGCTTTTTT 554
DB 811 TGAAGAGGAAATGCTTCTTACAAATATGCTCTCTCTGGTGGTCTCTCTATGGAATTTT 870
QY 555 AGCATATGCCCAATTTAAACGACACAGAAAGTCCCACTTATTGAATATCTCTTCTCTCTCTGA 614
DB 871 GGCATATGCGGAATTTAATGATAAGGATATACCAAGTATTTCCTTATTTTCAAGTGGCT 930
QY 615 TGAGTCACTGCGAGTACCCAAAGACTGTACGGGTTTCCATATCCAAAGCGAGGCTGTGAA 674
DB 931 TGA-----ACAATATCTAGAACAAATAAATTTTCCATATACCCAAAGGCTGGAGCTTAAAGAA 984
QY 675 TCCAACTGTAAAGTTCTTTTGTAAATACAGACTCTCTCAGCTCAGTCAACAATGCAAC 734
DB 985 TCCCGTGTTCGGATATTTATATCGATAC-----CACTTACCTGGGTATGTAGG 1035
QY 735 TTCCATACAAATCACTGCTCTCTCTCTTATGTTGATGGGATCACTACTTGTGTGATGT 794
DB 1036 TCCCGCAGGAGTGGCTGTTCAGCAATGATAGCCCTCAAGTCATTTATTTTCAAGTGGCT 1095
QY 795 GACATGGGCAACACAGAAAGAAATTTCTTTGCGTGGCTCAGGAGGATTCAGAGACTATTC 854
DB 1096 CACGTGGGTACTGATGAACAGGATGTTTGTGCGTGGCTAAAAAGAGTCCAGAAATGTTTC 1155
QY 855 GGTCACTGATATTTGTGACTATGATGAATCCAGTGGAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 914
DB 1156 GGTCTCTCTATATGCTACTTCAGGGAAGACTGCGAGACATGGGATTTGTCCAAAGACCA 1215
QY 915 GCACACATTTGAAATGAGTACTACTGCTGGGTTGGAAGATTTTAGGCTTTCAGAACTTCA 974
```

Db 1216 GGAGCATATAGAGAAGAGCAGAACTGGATGGGCTGGTGGATTCTTTGTTTCAAGCAAGT 1275  
 Qy 975 TTTTACCCCTTGATGGTAATAGCTTCTACAGATCATCAGCAATGAAGAAGGTTACAGACA 1034  
 Db 1276 TTTTACGCTATGATGCCATTTCGTACTACAAAATATTAGTGCAAGGATGGCTACAAACA 1335  
 Qy 1035 CATTTCTCTATTTCCAAATAGATAAAAAGACTGCACATTTATTACAAAAGGCACCTGGGA 1094  
 Db 1336 TATTCACTATATCAAGACACACTGTGGAAATGCTATTCAAATTCAGTGGCAAGTGGGA 1395  
 Qy 1095 AGTCATCGGATAGAGAAGCTCTAACAGTGTATATCTATACATTAATAGTAATGAATATAA 1154  
 Db 1396 GGCCATAAAATATATTACAGAGTAACACAGGATTCACCTGTTTATTCTAGCAATGAATTTGA 1455  
 Qy 1155 AGGAATGCCAGGAGGAGGATCTTTTATAAATCCAACTTATTGACTA---TACAAAAGT 1211  
 Db 1456 AGAATACCCCTGGAGAGAAACATCTACAGAAATAGCATTTGGAAGCTATCTCCCAAGCAA 1515  
 Qy 1212 GACATGCCCTCAGTTGTGAGCTGAATCCGGAAGGTGTGAGTACTATTCTGTGTCAATTGAG 1271  
 Db 1516 GAAGTGTGTACTTGGCATCTAAGGAAGAAAGGTGCCAATATTACACAGCAAGTTTCAG 1575  
 Qy 1272 TAAAGAGGCAAGTATTATCAGCTGAGATGTTCCGGTCTGGTCTGCCCTCTATACTCT 1331  
 Db 1576 CGACTACGCCAAGTACTATGCACTTGTCTGCTACGCCCCAGGCATCCCCATTTCCACCT 1635  
 Qy 1332 ACACAGCAGGCTGAATGATTAAGGCTGAGAGTCCCTGGAGACAATTCAGCTTTGGATAA 1391  
 Db 1636 TCATGATGACGCACTGATCAAGAAATTAATAATCTCTGSAAGAAACAAAGGAATTTGAAA 1695  
 Qy 1392 AATGCTGCAGATGTCCAGATGCCCTCCAAAACCTGGACTTCATTATTTTGAATGAAAC 1451  
 Db 1696 TGCCTTTGAANAATATCCAGCTGCTTAAGAGGAAATTAAGAACTTGAAGTAGATGAAT 1755  
 Qy 1452 AAAAATTTGGTATCAGATGATCTTGCTCTCTCAATTTGATAAATCAAGAAATATCCTCT 1511  
 Db 1756 TACTTTATGGTACAAGATGATTTCTCTCTCAATTTGACAGATCAAGAAGTATCCCTT 1815  
 Qy 1512 ACTATTAGATGTATGCHAGGCCCATGTAGTCAAAAGACAGACATGCTTTCAGACTGAA 1571  
 Db 1816 GCTAATTTCAAGTGTATGGTGGTCCCTGCAGTCAGAGTCTGATTTGCTGTTAA 1875  
 Qy 1572 CTGGGCCACTTACTTGCAGACACAGAAACATTTATAGTAGCTAGCTTTGATGGCAGAGG 1631  
 Db 1876 TTGGATATCTTATCTTGAAGTAAGAGAGGATGGTCTATGCTTGGTGGATGGTGGAG 1935  
 Qy 1632 AAGTGGTTTCAAGAGATGAAGATCATGATGCAATCAACAGAGACTGGGAACATTTGA 1691  
 Db 1936 AACAGCTTTCCAAGGTGACAAACTCTCTATGCAAGTGTATCGAAAGCTGGGTGTTATGA 1995  
 Qy 1692 AGTTGAAGATCAAAATGAAGCAGCCAGACAAATTTTCAAAATGGGATTTGTGACAAACA 1751  
 Db 1996 AGTTGAAGACCAAGATTACAGCTGTGAGAAATTTATAGAAATGGGTTCATTGATGA 2055  
 Qy 1752 ACGAATTTGCAATTTGGGGCTGGTCATATGAGGGTACGTAACTCAATGCTCTGGGATC 1811  
 Db 2056 AAGAAATGCCATATGAGGGCTGGTCTATGAGAGATAGCTTTCACTGCGCCCTTGATC 2115  
 Qy 1812 GGAAGTGGCGTGTTCAGATGTGGAATAGCCGCTGGCGCTGTATCCCGTGGGAGTACTA 1871  
 Db 2116 TGGAACTGGTCTTTTCAAAATGGTATAGCAGTGGCTCCAGTCTCCAGCTGGGAATATTA 2175  
 Qy 1872 TGACTCAGTGTACACAGAAGTTATATGGGTCTCCCACTCCAGTCTCCAGTGGGAATATTA 1931  
 Db 2176 CGCGTCTGTCTACACAGAGAGATTCATGGGTCTCCCAACAAAGGATGATAATCTTGAGCA 2235  
 Qy 1932 TTACAGAAATTCACAGTATGAGCAGAGCTGAAATTTTAAACAGATTTGAGTACCTCT 1991  
 Db 2236 CTATAGAATTTCACTGTGATGGCAAGACAGAAATTTTCAGAAATGTAGATATCTTCT 2295  
 Qy 1992 TATTCAAGAGCAGCAGATGATAACGTTCACTTTACAGAGTCAAGCTCAGATCTCCAAAGC 2051  
 Db 2296 CATCCAGGAAACAGCAGATGATAATGTGCACTTTCAAACTTCAGCACAGATTCGCTAAAGC 2355

Qy 2052 COTGGTGCATGTTGGAGTGGATTTCCAGGCAATGTGGTATATCTGATGAAGACCATGGAAT 2111  
 Db 2356 TCTGGTTAATGCACAAGTGGATTTCCAGGCATGTGGTACTCTGACCAGAACCCAGGCTT 2415  
 Qy 2112 AGCTAGCAGCAGACACACCAT 2171  
 Db 2416 ATCCGGCTGTCCAGAAC---CACTTATACACCCACATGACCCACTTCTTAAGCAGTG 2472  
 Qy 2172 TTTCTCTTT 2180  
 Db 2473 TTTCTCTTT 2481

RESULT 8

US-08-619-280A-1  
 ; Sequence 1, Application US/08619280A  
 ; Patent No. 5767242  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Zimmermann, Rainer; Park, John E.;  
 ; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.  
 ; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN  
 ; TITLE OF INVENTION: ALPHA, AND USES THEREOF  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Felfe & Lynch  
 ; STREET: 805 Third Avenue  
 ; CITY: New York City  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10022  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage  
 ; COMPUTER: IBM PS/2  
 ; OPERATING SYSTEM: PC-DOS  
 ; SOFTWARE: Wordperfect  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/619,280A  
 ; FILING DATE: 18-MARCH-1996  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/230,491  
 ; FILING DATE: 20-APRIL-1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Hanson, No 5767242man D.  
 ; REGISTRATION NUMBER: 30,946  
 ; REFERENCE/DOCKET NUMBER: LUD 5330.1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 688-9200  
 ; TELEFAX: (212) 838-3884  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2815 Base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; US-08-619-280A-1

Query Match 34.4%; Score 751; DB 2; Length 2815;  
 Best Local Similarity 60.9%; Pred. No. 1e-218;  
 Matches 1320; Conservative 0; Mismatches 825; Indels 24; Gaps 5;

Qy 15 CACTCTAACTGATTACTTTAAAAAATACCTTATAGACTGAAGTTATCTCTTAAAGATGGAT 74  
 Db 334 CACACTGAAGGATATATTTAAATGGAAACATTTTCTTATAAAACATTTTTCCAAACCTGGAT 393  
 Qy 75 TTCACATCATGATATCTCTACAAACAAGAAATAATATCTTGGTATTCAATGCTGAATA 134  
 Db 394 TTCAGACAGAGATATCTTCATCATCTGCAGATACCATATAGTACTTTTATANTATTCGA 453  
 Qy 135 TGGAAACAGCTCAGTTTTCTTGGAGAACAGTACATTTGATGAGTTGGACATCTTATCAA 194  
 Db 454 AACAGGACAAATCATATACCATTTTGGATTAATAGAACCAATGAAAAAGTGTGAATGCTTCAA- 512

RESULT 9  
US-08-940-391-1  
; Sequence 1, Application US/08940391  
; Patent No. 5965373  
; GENERAL INFORMATION:  
; APPLICANT: Zimmermann, Rainer; Park, John E.;  
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.  
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION



TITLE OF INVENTION: PROTEIN ALPHA, AND USES THEREOF

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Felice & Lynch

STREET: 805 Third Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/940,391

FILING DATE: 01-OCT-1997

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/619,280

FILING DATE: 18-MARCH-1996

APPLICATION NUMBER: 08/230,491

FILING DATE: 20-APRIL-1994

ATTORNEY/AGENT INFORMATION:

NAME: Hanson, No. 5965373man D.

REGISTRATION NUMBER: 30,946

REFERENCE/DOCKET NUMBER: LUD 5330.1

TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2815 Base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-940-391-1

Query Match 34.4%; Score 751; DB 2; Length 2815;

Best Local Similarity 60.9%; Pred. No. 1e-218;

Matches 1320; Conservative 0; Mismatches 825; Indels 24; Gaps 5;

QY	15	CACCTCACTGATCTTCTTAAATAATCTTATGACCTGAAAGTTATATCTCTCTTAAGTGGAT	74
DB	334	CACACTGAAGGATATTTTAAATGGAACATTTCTTATAAACAATTTTTCCTTAACTGGAT	393
QY	75	TTCCAGATCATGAATATCTCTACAAAACAGAAAATAATATCTTGGTATTCATGCTGAATA	134
DB	394	TTCCAGGACAAGAATATCTTCAATCTGCAGATAACAATATAGTACTTTTATAATATGGA	453
QY	135	TGMAAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTTGGACATCTCTATCAA	194
DB	454	AACAGGACAATCATATACCAATTTTGAGTAATAGAACCATGAAAGTGTGAATGCTTCAA-	512
QY	195	TGATATTCAATATCTCTGATGGGAGTTTATCTCTTAGAATAACAATCTAGCTGAAGCA	254
DB	513	--ATTACGGCTTATCACTGATCGGCAATTTGATATCTAGAAAGTATTTTCAAGCT	570
QY	255	ATGGAGGCTTCCTACACAGCTTTCATATGACATTTATGATTTTAAATTTAAAGGAGCTGAT	314
DB	571	TTGGAGATATCTTTACACAGCAACATATTACATCTATGACCTTTAGCAATGGAGAAATTC	630
QY	315	TACAGAAGAGAGATTTCCAAAACACACAGCTGGGTCAATGGTCAACAGTGGGTCTATAA	374
DB	631	AAGAGGAAATGAGCTTCTCTCGTCCAAATTCAGTATTTTATGCTGGTCCCTGTTGGGAGTAA	690
QY	375	ATTGGCATATGTTTGGAAACAAAGACATTTATGTTTAAATTTGAACCAAAATTTTCAAGTTA	434
DB	691	ATTAGCATATGCTATCAAAACAATATCTATTTTGAACCAAGACAGGAGATCCACCTTT	750
QY	435	CAGAATCACATGACGGGGAAGAAGATATAATATATAATGGAATTAATGGAATTAATGCTGGT	494
DB	751	TCAATAACATTTAATGGAAGAGAAAATAAATAATTTAATGGAATCCAGACTGGGTTTA	810

QY	495	TGAAGAGGAAGTCTTTCAGTGCCTACTGCTCTGTGGTGGTCTCCAAACGGCACTTTTT	554
DB	811	TGAAGAGGAATGCTTCTTACAAAATATGCTCTCTGTGGTCTCTCTTAATGGAATAATTT	870
QY	555	AGCATATGCCCAATTTAAACGACACAGAAAGTCCCACTTATTTGAATACTCTCTTCTACTCTGA	614
DB	871	GGCATATGGGAATTTAATGATAAGGATATACCACTTATTCCTATTTATGCGCA	930
QY	615	TGAGTCACTGCAGTACCCCAAGACTGTACGGGTCCATATCCAAAGGCGAGGCTGTGAA	674
DB	931	TGA-----ACAATATCTTAGAACAAATAATATTCATACCCCAAAGGCTGGAGCTAAGAA	984
QY	675	TCCAACTGTAAAGTTCTTTTGTGTAAATACAGACTCTCTCAGCTCAGTCAACCAATCCAAC	734
DB	985	TCCCGTTGTCGGATATTTATTCGTATAC-----CACTTACCCTGGGTATGTAGG	1035
QY	735	TTCCATACAAAATCACTGCTCTGCTTCTATGTTGATAGGGGATCACTACTTGTGTGATGT	794
DB	1036	TCCCCAGGAAGTGCCTGTTCCAGCAATGATAGCTCAAGTGATTTATTTTTCAGTTGGCT	1095
QY	795	GACATGGGCAACACAAAGAAATTTCTTTCAGTGGCTCAGAGGATTCAGAACTATTC	854
DB	1096	CACGTGGGTACTGATGAACGAGTATGTTTGCAGTGGCTTAAAGAGAGTCCAGAACTTTTC	1155
QY	855	GGTCATGGATATTTTGTGACTATGATGAATCCAGTGGAAAGATGGAACCTGCTTAGTGCAAG	914
DB	1156	GGTCTCTGTATATGCTGACTTCAGGGAAGACTGGCAGACATGGATTTGCCAAAGACCCA	1215
QY	915	GCAACACATTTGAAATGAGTACTACTGCTGGCTGGGTTGGAGATTTAGGCTTTCAGAACTCA	974
DB	1216	GGAGCATATAGAAGAAAGCAGAACTGGATGGCTGGTGGATTTCTTTGTTTCAAGACCAGT	1275
QY	975	TTTTTACCCTTGATGTAATAGCTTCTACAAGATCATCAGCAATGAAGAGGTTTACAGACA	1034
DB	1276	TTTCAGCTATGATGCCATTTTCGTACTACAAAATATTTAGTCAAGGATGGCTTACAAAACA	1335
QY	1035	CATTTGCTATTTTCCAAATAGATAAAAAAGACTGCACTTTATTAACAAAAGGACCTGGGA	1094
DB	1336	TATTCATATATCAAAAGACACTGTGGAAAATGCTATTCAAAATTAACAGTGGCAAGTGGGA	1395
QY	1095	AGTCATCGGGATPAGAAGCTCTAAACAGTGAATTTATCTACTACTATTAAGTAAATATAA	1154
DB	1396	GGCCATAAATAATTTTCAGAGTAAACACAGGATTCACCTGTTTATTTCTAGCAATGAATTTGA	1455
QY	1155	AGGAATGCCAGAGAGAGAAATCTTTATAAATCCAACTTATTTGACTA---TACAAAAGT	1211
DB	1456	AGAAATACCTCGGAAGAAGAAACATCTACAGAAATTAGCATTTGAAAGCTATCTCCCAAGCAA	1515
QY	1212	GACATGCCCTCAGTTGTGAGCTGAATCCGGAAGGTGTGAGTACTATTCTGTGTCTATTCTAG	1271
DB	1516	GAAGTGTGTTACTTGGCATCTAAGGAAGAAAGGTGCCAATATTACACAGCAAGTTTCAG	1575
QY	1272	TAAAGAGGCGAAGTATTTATTCAGCTGAGATGTTCCGGTCTGGTCTGCCCCCTCTATACTCT	1331
DB	1576	CGACTACGCCAAGTACTATGCACTTGTCTGCTACGGCCAGGCAATCCCCATTTCCACCCT	1635
QY	1332	ACACAGCAGCGTAATGATAAAGGGCTGAGAGTCTCTGGAGACAAATTCAGCTTTGGATAA	1391
DB	1636	TCATGATGGACGCATGATCAAGAAATTTAAAAATCTCTGGGAAGAAACAAAGAAATTTGAAA	1695
QY	1392	AATGCTGCAAGATGTCAGATGCCCTCCAAAAAATCTGGAATTTTCAATTTTGAATGAAAC	1451
DB	1696	TGCTTTGAAAATATCCAGCTGCCCTAAGAGGAAATTAAGAAACTTGAAGTAGATGAAAT	1755
QY	1452	AAAAATTTGGTATCAGATGATCTTCCCTCTCTTATTTGATAAATCCAAAGAAATATCTCTCT	1511
DB	1756	TACTTTATGGTCAAGATGATTTCT	1815
QY	1512	ACTATTAGATGATGACAGGCCCATGTAGTCAAAAAGCAGACACTGCTCTTCAGACTGAA	1571
DB	1816	GCTAATTTCAAGTGTATGGTGGTCCCTGCAAGTCAAGGTCTGTATTTTCTGCTGTAA	1875

```

QY 1572 CTGGCCACTTACCTTGAGCAGCAGCAAGAAACATTAATAGTAGCTAGCTTTGATGGCAGAGG 1631
Db 1876 TTGGATATCTTATCTTGAAGTAAGGAGGATGCTCATTTGCTTGGTGGATGGTCGAGG 1935
QY 1632 AAGTGGTTACCAAGGAGATAGATCATGCTGATGATGATGATGATGATGATGATGATGATGAT 1691
Db 1936 AACAGCTTTTCAAGGTGCAAACTCCTTATGATGATGATGATGATGATGATGATGATGATGAT 1995
QY 1692 AGTTGAAGATCAAAATTTGAAGCAGCAGCAGCAAAATTTTCAAAATTTGATGATGATGATGAT 1751
Db 1996 AGTTGAAGCAGATTAACGCTGTGCAAAATTTTCAAAATTTGATGATGATGATGATGATGAT 2055
QY 1752 AGCAATGCAATTTGGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1811
Db 2056 AAGAATAGCATTATGGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2115
QY 1812 GGGAGTGGCTGTTCAAGTGTGGATAGCCGCTGGGCTGTATCCCGTGGGAGTACTA 1871
Db 2116 TGGAACTGGTCTTTTCAAAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 2175
QY 1872 TGACTCAGTGTACAGCAAGCTTTACATGGGCTTCCCAACTCCAGAGCAACACCTTGACCA 1931
Db 2176 CGCGTCTCTTACAGCAGAGATTTCAATGGGCTTCCCAACAAAGATGATGATGATGATGAT 2235
QY 1932 TTACAGAAATCAACAGTCAAGCAGCAGCAGCTGATGATGATGATGATGATGATGATGATGAT 1991
Db 2236 CTATAAGAAATCAACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2295
QY 1992 TATTCAATGCAAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2051
Db 2296 CATCCAGCAAGCAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2355
QY 2052 CCTGTGCTGATTTGGAGTGGATTTCCAGGCAATGATGATGATGATGATGATGATGATGATGAT 2111
Db 2356 TCTGTTAATGCAAGTGGATTTCCAGGCAATGATGATGATGATGATGATGATGATGATGATGAT 2415
QY 2112 AGTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2171
Db 2416 ATCCGGCTGTCCAGAAC---CACTTATACCCACATGATGATGATGATGATGATGATGATGAT 2472
QY 2172 TTTCTCTTT 2180
Db 2473 TTTCTCTTT 2481

```

RESULT 10

```

US-09-265-606-1
; Sequence 1, Application US/09265606
; Patent No. 6846910
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John B.;
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
; TITLE OF INVENTION: ALPHA, AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felife & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09265,606
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,280

```

```

; FILING DATE: 18-MARCH-1996
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6846910man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2815 Base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-09-265-606-1

Query Match 34.4%; Score 751; DB 3; Length 2815;
Best Local Similarity 60.9%; Pred. No. 1e-218;
Matches 1320; Conservative 0; Mismatches 825; Indels 24; Gaps 5;

QY 15 CACTCTAACTGATTAATTAATAATACTTATAGACTGAGTATTAATCTCTTAAGATGAT 74
Db 334 CACACTGAAGGATATTTAAATGGAACATTTTCTTATAAAACATTTTTCCAAACCTGGAT 393
QY 75 TTCAGATCATCAATATCTTACAAACAAAGAAATAATCTTGGTATTCAATGCTGAATA 134
Db 394 TTCAGGACAGAAATATCTTCAATCTGAGATTAACAAATATAGTACTTTTATAATATGGA 453
QY 135 TGGAAACAGCTCAGCTTTTCTTGGAGAACAGTACATTTTGATGAGTTTGGACATTTCTATCAA 194
Db 454 AACAGGCAATCAATATACCATTTTGGAGTAAATAGAACATGAAGTGTGAATGCTTCAA- 512
QY 195 TGATTTATCAATATCTCTGATGGCAGTATTTATCTCTTGAATACAACTAGTGAAGCA 254
Db 513 --ATTACGGCTTATCAGCTGATCGGCAATTTGTATATCTAGAAAGTGTATTTCAAAGCT 570
QY 255 ATGAGGAGCATTTCTTACACAGCTTCATATGACATTTATGATTTAAATAAAGCAGCTGAT 314
Db 571 TTGAGGATATCTTACACAGCAATATATCATCTATGACCTTAGCAATGAGAAATTTGT 630
QY 315 TACAGAAAGAGAGGATTTCCAAACAAACACACAGTGGGTCACTGGTCCACAGTGGGTCAVAA 374
Db 631 AAGAGGAAATGAGCTTCTCTGCTCAATTCAGTATTTATGCTGTGCTGCTGTTGGAGTAA 690
QY 375 ATTGCGCATATGTTTGGAAACAAATGACATTTATGTTAAATTTGAACCAAAATTTACCAAGTTA 434
Db 691 ATTAGCATATGCTTATCAAAACAAATATCTATTTGAAACAAAGACCCAGGAGATCCACCTTT 750
QY 435 CAGAAATCAATGAGCGGGGAAAGAGATATAATATAATTAATGAAATTAATGAAATTTCCAGACTGGGTTTA 494
Db 751 TCAAAATCAATTTAATGGAAGAGAAATAATAATTTAATGAAATTTCCAGACTGGGTTTA 810
QY 495 TGAAGAGGAAAGTCTTTCAGTGCCTACTCTGCTCTGCTGGTGGTCTCTCCAAACGGCAGCTTTT 554
Db 811 TGAAGAGGAAATGCTTCTTCAAAATATGCTCTCTGCTGGTGTCTCTTAATGGAATTTT 870
QY 555 AGCATATGCCCAATTTAAACGACACAGAGTCCCACTTATGGAATATCTCTTCTACTCTGA 614
Db 871 GGCATATGCGGAATTTAATGATAAGGATATACCAAGTTATTCCTATTTATGCGGA 930
QY 615 TGAGTCACTGAGTACCCAAAGAGCTGTACGGGTTCCATATCCAAAGGAGGAGCTGTGAA 674
Db 931 TGA-----ACAATATCTTGAACAAATAATTTCCATATCCAAAGGCTGGAGCTTAAGAA 984
QY 675 TCCAACTGTAAAGTCTTTTGTGTTAAATACAGACTCTCTCAGCTCAGTCACTCAATGCAAC 734
Db 985 TCCGTTGTTCCGATATTTTATTCGATAC-----CACTTACCTGCGTATGATGAG 1035
QY 735 TTCCATACAAATCATCTGCTCTGCTCTTATGTTGATAGGGGATCACTACTTGTGTGATGT 794
Db 1036 TCCCAGGAAAGTGCCTGTTCAGCAATGATAGCCTCAAGCTCAAGTGTATTTATTTTCACTTGGCT 1095

```



Qy	315	TACAGAGAGAGAGATTCCAAACAACAACACACAGTGGGTACATGGTCCACAGTGGGTGATTA	374
Db	631	AAGAGGAATGAGCTTCCTCGTCCAAATTCAGTATTTATGCTGGTGGCTGTTGGAGGTAA	690
Qy	375	ATTGGCATATGTTTGGAAACAATGACATTTATGTTTAAATTTGAAATTTGAAATTTACCAAGTTA	434
Db	691	ATTAGCATATGCTCTATCAAAACATATCTATTTGAAACAAGACCGAGAGATCCACCTTT	750
Qy	435	CAGAAATCACATGGACGGGGAAGAAGATATAATATATAATATGGAATAACTGACTGGGTTTA	494
Db	751	TCAAATAACATTTAATGGAAGAGAAAAATAAATAATTTAAATGGAATCCAGACTGGGTTTA	810
Qy	495	TGAAGAGGAGTCTTTCAGTGGCTACTCTGCTCTGCTGGTGGTCTCCAAACGGCACTTTTTT	554
Db	811	TGAAGAGGAATGCTTGCTACAAATATGCTCTCTGCTGGTGGTCTCCTAATGGAAAAATTTTT	870
Qy	555	AGCATATGCCCAATTTAAACGACACAGAAAGTCCCACTTTATTGAAATACTCCTTCTACTCTGA	614
Db	871	GGCATATGGCGAATTTAATGATACGGATATACAGATTATTGGCTATTCTTATATGGCGA	930
Qy	615	TGAGTCACTGCAGTACCCAAAGACTGTACGGGTTCCATATCCAAAGGACGAGCTGTGAA	674
Db	931	TGAA-----CAATATCCTAGAACAAATAAATTTCCATACCCAAAGGCTGGAGCTAAGAA	984
Qy	675	TCCAACTGTAAAGTCTTTGTTGTTAAATACAGACTCTCTCAGCTCAGTCACCAATGCAAC	734
Db	985	TCCCGTTGTTGGGATATTTATATCGATAC-----CACTTACCCTCGCTATGTAGG	1035
Qy	735	TTCCATACAAATCACTGCTCTGCTCTTATGTTGATAGGGGATCACTACTTGTGTGATGT	794
Db	1036	TCCCCAGGAGTGCTGTTCCAGCAATGATAGCCTCAAGTGATTTATTTCAGTTGGCT	1095
Qy	795	GACATGGGCAACAAGAAAGAAATTTCTTTGCACTGGCTCAGGAGGATTCAGAACTATTC	854
Db	1096	CACGTGGGTTACTGATGAACGAGTATGTTTGCAGTGGCTAAAAGAGTCCAGAATGTTTC	1155
Qy	855	GGTCATGGATTTTGCACTATGATGAATCCAGTGGAGAGATGGAACTGCTTACTGGCAAG	914
Db	1156	GGTCCTGTCTATATGTGACTTCAGGGAAGACTGGCAGACATGGGATTTGCCAAGACCCA	1215
Qy	915	GCACACATTTGAAATGAGTACTACTGGCTGGGTTGGAGAGATTTTAGGCGTTTCAGAACCTCA	974
Db	1216	GGACATATAGAAGAAAGCAGACTGATGGCTGGTGGATCTTTGTTTCAACACCAGT	1275
Qy	975	TTTTACCTCTGATGGTAAATAGTCTTACAAGATCATCAGCAATGAGAAGAGTTACAGACA	1034
Db	1276	TTTCAGCTATGATGCCATTTTCGTACTACAAAATATTTAGTGACAAGGATGGCTACAAACA	1335
Qy	1035	CATTTGCTATTTCCAAATAGATAAAAAGACTGCACATTTATTACAAAAGGCACTGGGA	1094
Db	1336	TATTCACATATCAAGACACTGGGAAAATGCTATTCAAATTAACAAGTGGCAAGTGGGA	1395
Qy	1095	AGTCATCGGATAGAAGCTCTTAACACAGTGAATTTACTACTACATTAGTAAATGAATATA	1154
Db	1396	GGCCATAATATATTACAGATTAACACAGGATTCACGTGTTTATTCTAGCAATGAAATTGA	1455
Qy	1155	AGGATGCCAGGAGGAAGGAATCTTTTATAAAATCCAACTTATTAGACTATA---CAAAGT	1211
Db	1456	AGAAATACCTCGGAGAGAAAACTACAGAAATTAGCAATTTGGAAGCTACTCTCCAAGCAA	1515
Qy	1212	GACATGCTCAGTTGTGAGCTGGAATCCGGAAGAGGTGCAGTACTATTCTGTGTCATTCAG	1271
Db	1516	GAAGTGTGTACTTGGCCATCTAAGGAAGAAGAGGTGCCAATATTACACGAAGTTTCAG	1575
Qy	1272	TAAAGAGCGGAAGTATTATCAGCTGAGATGTTTCGGTCTGGTCTGCCCTCTATACTCT	1331
Db	1576	CGACTACGCCAAGTACTATGCATCTGTCTGT-----	1607
Qy	1332	ACACAGCAGCGTGAATGATTAAGGGCTGAGAGTCTCTGGAAAGACAATTCAGCTTTGGATAA	1391
Db	1608	-----ACGAAATTAATAATCTCTGGAAGAAAAACAAGGAATTTGGAATA	1647

QY	1392	AATGCTGCAGNATGTCAGATGCGCTCCAAAANAACCTGGACCTTCATTAATTTTGTATGAAC	1451
Db	1648	TGCTTTGAAAANAATATCAGCTGCCCTAAAGAGGAAAATTAAGAAACTTGAAGTAGATGAAT	1707
QY	1452	AAAAATTTTGGTATCAGATGATCTTGGCCCTCCTCAATTTTGTATAAATCCAAAGAAAATATCCTCT	1511
Db	1708	TACTTTATGGTACAGATGATCTTCTCTCTCAATTTTGAAGATCAAGAGATATCCTT	1767
QY	1512	ACTATTAGATGTGATGCA---GGCCCATGTGTAGTCAAAAAGCAGACACATGTCTTTCAGAC	1567
Db	1768	GCTAATTAACAGTGTATGTATGTTGGTCCCTGCGAGTCAGAGTGTAAAGGTCTGTATTTCCTG	1827
QY	1568	TGAAGTGGGCACTTACCTTGCAAGCAGACAGAAAACATATATAGTAGCTAGCTTTGTATGGCA	1627
Db	1828	TTAATTTGGATATCTTATCTTTGCAAGTAAAGAGGGATGGTCATTTGCCCTTCGGTGGATGGTC	1887
QY	1628	GAGGAAGTGGTTTACCAAGGAGATAAGATCATGCAATGCAATCAACAGAGAAGACTTGGGAACAT	1687
Db	1888	GAGGACAGCTTTTCCAAGGTGACAACTCTCTATGCAGTGTATCGAAGCTGGGTGTTT	1947
QY	1688	TTGAAGTTGAAGATCAAAATTTGAAGCAGCCAGACAAATTTTCAAAAATGGGAATTTGTGGACA	1747
Db	1948	ATGAAGTTGAAGACCAATTAACAGCTGTACAGAAAATTCATAGAAAATGGGTTTTCATTGATG	2007
QY	1748	ACAAACGAATTCGAATTTGGGGCTGGTCAATATGAGAGGTTACGTAAACCTCAATGGTCTCTG	1807
Db	2008	AAAAAAGAAATAGCCATATAGGGGCTGGTCTTATGAGAGGATACGTTTTCATCACTGGCCCTTG	2067
QY	1808	GATCGGGAAGTGGGTGTTCAAGTGTGGATAGCCGTGGCGCTGTATCCCGTGGGAGT	1867
Db	2068	CATCTGGAACCTGGTCTTTTCAAAATGTGGTATAGCAGTGGCTCCAGTCTCCAGCTGGGAAT	2127
QY	1868	ACTATGACTCAGTGTACACAGAACTGTACATGGGTCTCCCAAATCCAGAGAACAACCTTG	1927
Db	2128	ATTACGGCTGTGTCTACAGAGAGATTCATGGGTCTCCCAAAGAGATGATATCTTG	2187
QY	1928	ACCAATACAGAAATTCACAGTCAATGAGCAGAGCTGAAATTTTAAACAAAGTTGAGTACC	1987
Db	2188	AGCACTATAAGAAATTCAACTGTGATGCAAGAGAGCAGAAATATTTTCAGAAAATGTAGACTATC	2247
QY	1988	TCCTTATTCATGNACAGCAGATGATAACGTTCACTTTTCAGCAGTCAGCTCAGATCTCCA	2047
Db	2248	TTCTCATCCACGGAACAGCAGATGATAATGTGCACTTTTCAAAACTCAGCACAGATTGCTA	2307
QY	2048	AAGCCCTGGTTCGATGTTGGAGTGGATTTCCAGGCAATGTGGTATCTGATGAAGACCATG	2107
Db	2308	AAGCTCTGGTTATGACACAGTGGATTTCCAGGCAATGTGGTACTCTGACCAGAACCAAG	2367
QY	2108	GAATAGCTAGCAGCAGCAGCACCAACATATATATACCCAATGAGGCCACTTCATAAAAAC	2167
Db	2368	GCTTATCGGCTGTCCACGAAC---CACTTATACCCCAACATGACCCACTTCTCTAAAGC	2424
QY	2168	AATGTTTCTCTTT	2180
Db	2425	AGTGTTCCTCTTT	2437

RESULT 12

US-09-389-681-428

; Sequence 428, Application US/09389681A

; Patent No. 6518237

; GENERAL INFORMATION:

; APPLICANT: Yucui, Jiang

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Xu, Jiangchun

; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND

; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE

; FILE REFERENCE: 210121-470C3

; CURRENT APPLICATION NUMBER: US/09/389,681A

; CURRENT FILING DATE: 1999-09-02

; NUMBER OF SEQ ID NOS: 463

; SOFTWARE: Fast-SEQ for Windows Version 3.0

## RESULT 12

```

US-09-389-681-428
; Sequence 428, Application US/09389681A
; Patent No. 6518237
; GENERAL INFORMATION:
; APPLICANT: Yuqiu, Jiang
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.470C3
; CURRENT APPLICATION NUMBER: US/09/389,681A
; CURRENT FILING DATE: 1999-09-02
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0

```



```
OY 1586 TTGCAAGCAGACAGAAACATTATAGTAGCTAGCTTTGATGGCAGAGAAAGTGGTTACCAAG 1645
    |||||
Db 124 TTGCAAGTAAGGAAGGATGGTCATTGCTTGGTGGATGGTTCGAGGAACAGCTTTCCCAAG 183

OY 1646 GAGATAAGATCATGATGCAATCAACAGAGACTGGGAACATTGGAAGTTGAAGATCAAA 1705
    |||||
Db 184 GTGACAAACTCCTCTATGCGAGTGTATCGAAAGCTGGGTGTTTATGAAGTTGAAGACCAGA 243

OY 1706 TTGAAGCAGCAGACAAATTTTCAAAAATGGGATTTGTGGACAAACAAAGAAATTGCAATTT 1765
    |||||
Db 244 TTACAGCTGTGAGAAATTCATGAAATGGGTTTCATTGATGAAAAAAGAAATAGCCATAT 303

OY 1766 GGGCTGTGTCATATAGGAGGTACGTAACCTCAATGGTCTCTGGGATCGGGAAGTGGCGTGT 1825
    |||||
Db 304 GGGGCTGGTCTCTATGGAGATACGTTTCATCACTGGCCCTTGCATCTGGAACCTGCTCTTT 363

OY 1826 TCAAGTGTGGAATAGCCGTGGGCTGTATCCCGTGGGAGTACTATGACTCAGTGTACA 1885
    |||||
Db 364 TCAATGTGGTATAGCAGTGGCTCCAGTCTCCAGCTGGGAATATTTACCGCTCTGTCTACA 423

OY 1886 CAGAACGTTTACATGGGTCTCCCAACTCCAGAAAGACAACCTTGACCATTACAGAAATTCAA 1945
    |||||
Db 424 CAGAGAGATTTCATGGGTCTCCCAACAAAGGATGATATCTTGAGCACTATAGAAATTCAA 483

OY 1946 CAGTCATGAGCAGAGCTGAAATTTTAAACAAGTTGAGTACTCTCTTATTTCA 1997
    |||||
Db 484 CTGTGATGGCAAGACGACAAATATTTTCAGAAATGTAGACTATCTTCTCATCCA 535
```

RESULT 15  
US-09-604-287A-428  
; Sequence 428, Application US/09604287A  
; Patent No. 6586572  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yuguu  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Harlocker, Susan L.  
; APPLICANT: Hepler, William T.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE OF INVENTION: DIAGNOSIS OF BREAST CANCER  
; FILE REFERENCE: 210121.470C7  
; CURRENT APPLICATION NUMBER: US/09/604,287A  
; CURRENT FILING DATE: 2000-06-22  
; NUMBER OF SEQ ID NOS: 489  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 428  
; LENGTH: 535  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-604-287A-428

```
Query Match 12.1%; Score 264.8; DB 3; Length 535;  
Best Local Similarity 68.6%; Pred. No. 2.3e-70;  
Matches 365; Conservative 0; Mismatches 167; Indels 0; Gaps 0;  
  
OY 1466 AGATGATCTTGCCTCTCTCATTTTGATTAATCCAGAAATATCCCTCTACTATTAGATGTGT 1525  
    |||||  
Db 4 AGATGATCTTCTCTCTCTCAATTTGACAGATCAAAAGAGATATCCCTTGTCTAATTCAGTGT 63  
  
OY 1526 ATGCAGGCCCATGTAGTCAAAAAGCAGACACTGTCTTCAGACTGAACTGGGCCACTTACC 1585  
    |||||  
Db 64 ATGGTGTCTCTGAGTCAGAGTGTAGGTCTGTATTGCTGTTAATTTGGATATCTTATC 123  
  
OY 1586 TTGCAAGCAGACAGAAACATTATAGTAGCTTTTGTATGGCAGAGGAAGTGGTTACCAAG 1645  
    |||||  
Db 124 TTGCAAGTAAGGAAGGATGGTCATTGCTTGGTGGATGGTTCGAGGAACAGCTTTTCCCAAG 183  
  
OY 1646 GAGATAAGATCATGATGCAATCAACAGAGACTGGGAACATTTGAAAGTTGAAGATCAAA 1705  
    |||||  
Db 184 GTGACAAACTCCTCTATGCGAGTGTATCGAAAGCTGGGTGTTTATGAAGTTGAAGACCAGA 243
```

Search completed: February 17, 2006, 07:48:59  
Job time : 406 secs



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2006, 04:14:35 ; Search time 1316 Seconds  
(without alignments)  
11060.554 Million cell updates/sec

Title: US-10-659-055-2  
Perfect score: 2184  
Sequence: 1 agtcgcaaaactacactct.....aacaatgtttctcttacct 2184

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues  
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : N\_Geneseq\_21.\*
- 1: Geneseqn1980s.\*
  - 2: Geneseqn1990s.\*
  - 3: Geneseqn2000s.\*
  - 4: Geneseqn2001as.\*
  - 5: Geneseqn2001bs.\*
  - 6: Geneseqn2002as.\*
  - 7: Geneseqn2002bs.\*
  - 8: Geneseqn2003as.\*
  - 9: Geneseqn2003bs.\*
  - 10: Geneseqn2003cs.\*
  - 11: Geneseqn2003ds.\*
  - 12: Geneseqn2004as.\*
  - 13: Geneseqn2004bs.\*
  - 14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2184	100.0	2211	12	ADO40239 Human DPP
2	2184	100.0	3407	6	Abk92227 Prostate
3	2184	100.0	3407	6	Aal44351 Human dip
4	2184	100.0	3407	8	Abz22557 Human liv
5	2184	100.0	3407	10	Add14640 Human src
6	2184	100.0	3407	11	Adn39603 Cancer/an
7	2184	100.0	3407	12	Ado19399 Human PRO
8	2184	100.0	3407	13	Adv25524 Human dip
9	2182.4	99.9	2301	11	Adn39271 Cancer/an
10	2182.4	99.9	2301	12	Adj83980 Human ful
11	2182.4	99.9	2301	14	Adv43983 Human psy
12	2182.4	99.9	2461	12	Ado71643 Nucleotid
13	2182.4	99.9	3445	12	Adj74790 Marker ge
14	2182.4	99.9	3445	12	Ado19397 Human PRO
15	2182.4	99.9	3445	12	Ado71613 Nucleotid
16	2182.4	99.9	3445	13	Acn37783 Tumour-as
17	2182.4	99.9	3445	13	Adp54457 Human PRO
18	2182.4	99.9	3445	13	Adu06201 Novel bro
19	2182.4	99.9	3445	14	Ady16579 DNA encod

20	2182.4	99.9	3445	14	ADY15160
21	2182.4	99.9	3445	14	ADZ14037 Human dip
22	2182.4	99.9	3913	14	ADZ48887 Insulin s
23	2182.4	99.9	3913	14	ABE94224 CD26/dipe
24	2180.8	99.9	2924	2	AAQ46089 Sequence
25	2180.8	99.9	2924	10	ADL24801 Intestina
26	2180.8	99.9	2924	11	ADI31700 Human cDN
27	2180.8	99.9	2924	12	ADO19805 Human PRO
28	2180.8	99.9	2924	12	ADO71611 Nucleotid
29	2180.8	99.9	2924	13	ADS83767 Human lym
30	2179.2	99.8	2924	2	AAQ63261
31	2177.6	99.7	2375	13	ADQ80241 Dipeptidy
32	2170.4	99.4	3410	12	ADO71641 Nucleotid
33	2170.4	99.4	3410	13	ACF87398 Human SIR
34	2170.4	99.4	3410	13	ADL24802 Intestina
35	1677.6	76.8	3316	12	ADJ75717 Marker ge
36	1677.6	76.8	3316	12	ADO71645 Nucleotid
37	1677.6	76.8	3482	14	ABE94225
38	1656.2	75.8	4835	6	ABK63663
39	1656.2	75.8	4835	8	ABZ22556
40	1656.2	75.8	4835	10	ADB58114
41	1656.2	75.8	4835	10	ADB52598
42	1656.2	75.8	4835	12	ADP72568
43	754.2	34.5	2366	6	ABZ35572
44	752.6	34.5	2310	14	ABE94221
45	752.6	34.5	2788	11	ADN95553

ALIGNMENTS

RESULT 1	
ADO40239	
ID	ADO40239 standard; cDNA; 2211 BP.
XX	
AC	ADO40239;
XX	
DT	12-AUG-2004 (first entry)
XX	
DE	Human DPP-IV extracellular domain encoding cDNA SEQ ID NO:1.
XX	
KW	crystal; mammalian dipeptidyl-peptidase IV extracellular domain;
KW	dipeptidyl-peptidase IV extracellular domain;
KW	DPP-IV extracellular domain; three-dimensional structure; antidiabetic;
KW	anorectic; cytostatic; type I diabetes; type II diabetes; IGT; obesity;
KW	cancer; human; DPP-IV; enzyme; gene; ss; protein co-ordinate data;
KW	EC 3.4.14.5.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	CDS 1..2211
FT	/*cag= a
FT	/product= "DPP-IV extracellular domain"
FT	/EC_number= "3.4.14.5"
XX	
PN	EPI422293-A1.
XX	
PD	26-MAY-2004.
XX	
PP	17-NOV-2003; 2003EP-00026169.
XX	
PR	25-NOV-2002; 2002EP-00026367.
XX	
PA	(HOFF ) HOFFMANN LA ROCHE & CO AG F.
XX	
PI	Hennig M, Loeffler EM, Thoma R;
XX	
DR	WPI; 2004-413363/39.
XX	
DR	P-PSDB; ADO40240.
XX	
PT	New crystal of an extracellular domain of mammalian dipeptidyl-peptidase
PT	IV (DPP-IV) useful for identifying or designing inhibitors of DPP-IV

PT activity.

XX Claim 26; SEQ ID NO 1; 215pp; English.

XX The present invention describes a crystal (I) of the extracellular domain of mammalian dipeptidyl-peptidase (DPP)-IV (EC 3.4.14.5). Also described: (1) a co-crystal of the extracellular domain of mammalian DPP-IV and a ligand bound to its active site; (2) a co-crystal of the extracellular domain of mammalian DPP-IV and a ligand bound to an allosteric binding site; (3) a co-crystal of the extracellular domain of mammalian DPP-IV and HgCl<sub>2</sub>; (4) crystallising (M1) mammalian DPP-IV; (5) co-crystallising (M2) mammalian DPP-IV and an active site ligand; (6) a crystal produced by (M1) and (M2); (7) determining the three-dimensional structure of a crystallised extracellular domain of mammalian DPP-IV to a resolution of 3.5-2.1 angstrom or better; (8) a machine-readable data storage medium comprising a data storage material encoded with machine readable data which, when using a machine programmed with instructions for using the data, displays a graphical three-dimensional representation of a molecule or molecular complex comprising at least a portion of the extracellular domain of mammalian DPP-IV comprising a fully defined sequence (SEQ ID NO:2, SI) of 736 amino acids, where the extracellular domain comprising the ligand binding active site being defined by a set of points having a root mean square deviation of less than about 1.5 angstrom from points representing the backbone atoms of the amino acids as represented by structure coordinates as given in the specification; (9) a compound (II) identified by using (I); (10) a pharmaceutical composition (III) comprising (I) and a carrier; (11) an isolated nucleic acid sequence (IV) encoding the soluble extracellular domain of DPP-IV comprising a fully defined sequence (SEQ ID NO:1, S2) of 2211 nucleotides; (12) a nucleic acid construct (V) comprising an expression vector and (IV); (13) a host cell (VI) transformed with (V); (14) producing the soluble extracellular domain of DPP-IV, involves culturing (VI) under conditions permitting the expression of the soluble extracellular domain of DPP-IV by (VI); and (15) a polypeptide comprising the soluble extracellular domain of (SI). DPP-IV has antidiabetic, anorectic and cytostatic activities. (I) is useful for identifying a compound that interacts with DPP-IV. The compound interacts with the active site of DPP-IV. The compound is an inhibitor of DPP-IV activity. (I) is useful for the identification and/or design of inhibitors of DPP-IV activity. (II) is useful as a therapeutic active substance, in particular for the treatment of diabetes type I, diabetes type II, IGT, obesity and cancer. (II) is useful for the manufacture of a medicament for the treatment of above mentioned disease. The present sequence encodes the extracellular domain of human DPP-IV, which is used in the exemplification of the present invention.

XX SQ Sequence 2211 BP; 716 A; 429 C; 466 G; 600 T; 0 U; 0 Other;

Query Match 100.0%; Score 2184; DB 12; Length 2211;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTCGAAAACCTTACACTCTAAGTATGATCTTAAATAAATCTTATAGACTGAAGTTATAC 60  
DB 25 AGTCGAAAACCTTACACTCTAAGTATGATCTTAAATAAATCTTATAGACTGAAGTTATAC 84

QY 61 TCCTTAAGATGGATTTTCAGATCATGATATCTCTACAAAACAGAAAATATATCTGGTA 120  
DB 85 TCCTTAAGATGGATTTTCAGATCATGATATCTCTACAAAACAGAAAATATATCTGGTA 144

QY 121 TTCAATGCTGAAATATGGAACAGCTCAGTTTCTTCGGAGAACAGTACATTTGATGAGTTT 180  
DB 145 TTCAATGCTGAAATATGGAACAGCTCAGTTTCTTCGGAGAACAGTACATTTGATGAGTTT 204

QY 181 GGACATTTCTATCAATGATTTATTCATATATCTCTGATGGCAGGTTTATCTCTTAGAATAC 240  
DB 205 GGACATTTCTATCAATGATTTATTCATATATCTCTGATGGCAGGTTTATCTCTTAGAATAC 264

QY 241 AACTAGCTGAAGCAATGGAGGCAATTCCTACAGCTTCATATGACATTTATGATTTAAAT 300  
DB 265 AACTAGCTGAAGCAATGGAGGCAATTCCTACAGCTTCATATGACATTTATGATTTAAAT 324

QY 301 AAAAGSCAGCTGATTACAGAGAGAGGATTCCAAACAACACACAGTGGGTCACTGGTCA 360

DB 325 AAAAGSCAGCTGATTACAGAGAGAGGATTCCAAACAACACACAGTGGGTCACTGGTCA 384  
QY 361 CCAGTGGGTCAATAAATGGGCATATGTTTGGACAAATGACATTTATGTTAAAAATTCGAACA 420  
DB 385 CCAGTGGGTCAATAAATGGGCATATGTTTGGACAAATGACATTTATGTTAAAAATTCGAACA 444  
QY 421 AATTTACCAGATTTACAGAAATCACTGACGGGAAAGAGATATATATATATATGGAATA 480  
DB 445 AATTTACCAGATTTACAGAAATCACTGACGGGAAAGAGATATATATATATATGGAATA 504  
QY 481 ACTGACTGGGTTTATGAAGAGAGAGTCTTCAGTGCCTACTCTGCTCTCTGTGGTCTCA 540  
DB 505 ACTGACTGGGTTTATGAAGAGAGAGTCTTCAGTGCCTACTCTGCTCTCTGTGGTCTCA 564  
QY 541 AACGGCACTTTTATGACATATGCCCAATTTAACGACAGAGAGTCCCACTTATTCGAATAC 600  
DB 565 AACGGCACTTTTATGACATATGCCCAATTTAACGACAGAGAGTCCCACTTATTCGAATAC 624  
QY 601 TCCTTCTACTCTGATGAGTCACTGACAGTACCCAAAGACTGTACGGGTTCCATATCCAAAG 660  
DB 625 TCCTTCTACTCTGATGAGTCACTGACAGTACCCAAAGACTGTACGGGTTCCATATCCAAAG 684  
QY 661 GCAGAGCTGTGAATCCAACTGTAAGTCTTTTGTGTAATAACAGACTCTCTCAGCTCA 720  
DB 685 GCAGAGCTGTGAATCCAACTGTAAGTCTTTTGTGTAATAACAGACTCTCTCAGCTCA 744  
QY 721 GTCACCAATGCAACTTCCATACAACTACTGCTCTCTCTCTCTATGTTGATAGGGATCAC 780  
DB 745 GTCACCAATGCAACTTCCATACAACTACTGCTCTCTCTCTCTATGTTGATAGGGATCAC 804  
QY 781 TACTTGTGTGATGTGACATGGGCAACACAGAAAGAAATTTCTTTCAGTGGGTCCAGGAG 840  
DB 805 TACTTGTGTGATGTGACATGGGCAACACAGAAAGAAATTTCTTTCAGTGGGTCCAGGAG 864  
QY 841 ATTCAAGAACTATTCGGTCAATGGATTTGTGATATGATGAATCCAGTGGGAAGATGGAAC 900  
DB 865 ATTCAAGAACTATTCGGTCAATGGATTTGTGATATGATGAATCCAGTGGGAAGATGGAAC 924  
QY 901 TGCTTAGTGGCAGCGCAACACATTCGAATGATCTACTGCGTGGGTGGAGAGATTTAGG 960  
DB 925 TGCTTAGTGGCAGCGCAACACATTCGAATGATCTACTGCGTGGGTGGAGAGATTTAGG 984  
QY 961 CCTTCAGAAACCTCATTTTACCTTCCATGATAGCTTCTACAGATCATCAGCAATGAA 1020  
DB 985 CCTTCAGAAACCTCATTTTACCTTCCATGATAGCTTCTACAGATCATCAGCAATGAA 1044  
QY 1021 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATAAAAAAGACTGCACTTATATACA 1080  
DB 1045 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATAAAAAAGACTGCACTTATATACA 1104  
QY 1081 AAAGGCACCTGGGAATGTCATCGGATAGAGCTCTTAAACAGTGAATTTCTATCTACTACAT 1140  
DB 1105 AAAGGCACCTGGGAATGTCATCGGATAGAGCTCTTAAACAGTGAATTTCTATCTACTACAT 1164  
QY 1141 AGTAATGAATATAAAGGAATGCCAGGAGAGGAATCTTTATAAAATCCAACTTATTTGAC 1200  
DB 1165 AGTAATGAATATAAAGGAATGCCAGGAGAGGAATCTTTATAAAATCCAACTTATTTGAC 1224  
QY 1201 TATACAAAGAGTGCATGCTCAGTTGTGAGCTGAAATCCGGAAAGGTGTGAGTACTATTCT 1260  
DB 1225 TATACAAAGAGTGCATGCTCAGTTGTGAGCTGAAATCCGGAAAGGTGTGAGTACTATTCT 1284  
QY 1261 GTGTCATTTCAAGAGCGGAAGTATTTATTCAGCTGAGATGTTCCGGTCTCTGCTGCCCC 1320  
DB 1285 GTGTCATTTCAAGAGCGGAAGTATTTATTCAGCTGAGATGTTCCGGTCTCTGCTGCCCC 1344  
QY 1321 CTCCTACTCTACACAGCAGCGTGAATGATAAAGGCTCAGAGTCTCGGAAGACAATTCA 1380  
DB 1345 CTCCTACTCTACACAGCAGCGTGAATGATAAAGGCTCAGAGTCTCGGAAGACAATTCA 1404  
QY 1381 GCTTTGGATATAAATGCTCAGAAATGTCCAGATGTCCCTCCAAAAAAGTCTGAGTCTATT 1440



490	AAAAGGCAGCTGATTTACAGAAGAGAGGATTCCAAACAACACACAGTGGGTTCACATGGTCA	549
361	CCAGTGGGTCAATAAATTGGCATAATGTTTGGAAACAATGACATTTATGTATTAATAATTGAACCA	420
550	CCAGTGGGTCAATAAATTGGCATAATGTTTGGAAACAATGACATTTATGTATTAATAATTGAACCA	609
421	AATTTACCAAGTTACAGAAATCACATGGAACGGGGAAAGAGATATAAATATATATATGGAATA	480
610	AATTTACCAAGTTACAGAAATCACATGGAACGGGGAAAGAGATATAAATATATATGGAATA	669
481	ACTGACTGGGTATTGAAGAGGAAGTCTTCAGTGCCTACTCTGCTCTGTGGTGGTCTCCCA	540
670	ACTGACTGGGTATTGAAGAGGAAGTCTTCAGTGCCTACTCTGCTCTGTGGTGGTCTCCCA	729
541	AACGGGCACCTTTTATAGCATATGCCCAATTTAAACGACACAGAAGTCCCACTTTATTGAATAC	600
730	AACGGGCACCTTTTATAGCATATGCCCAATTTAAACGACACAGAAGTCCCACTTTATTGAATAC	789
601	TCCTTTCTACTCTGATGAGTCACTGCAGTACCCAAAGAATGTTACGGGTTCCATATPCCAAAG	660
790	TCCTTTCTACTCTGATGAGTCACTGCAGTACCCAAAGAATGTTACGGGTTCCATATPCCAAAG	849
661	GCAGGAGCTGTGAATCCAACTGTAAAGTCTTTTCTGTCTTAATACAGACTCTCTCAGCTCA	720
850	GCAAGGAGCTGTGAATCCAACTGTAAAGTCTTTTCTGTCTTAATACAGACTCTCTCAGCTCA	909
721	GTCAACCAATGCAACTTCATACAAATCACTGCTCCTGCTTCTATGTTGATAGGGGATCAC	780
910	GTCAACCAATGCAACTTCATACAAATCACTGCTCCTGCTTCTATGTTGATAGGGGATCAC	969
781	TACTTGTGTGATGTGACATGCGGCAACACAGAAGAAATTTCTTTGCAGTGGCTCAGGAGG	840
970	TACTTGTGTGATGTGACATGCGGCAACACAGAAGAAATTTCTTTGCAGTGGCTCAGGAGG	1029
841	ATTTCAGAACTATTCCGTTCATGGATATTGTGCACTATGATGAATCCAGTGGAGNATGCAAC	900
1030	ATTTCAGAACTATTCCGTTCATGGATATTGTGCACTATGATGAATCCAGTGGAGNATGCAAC	1089
901	TGCTTTAGTGGCACGGCAACACATTTGAAATGAGTACTACTGGCTGGGTGGAAAGATTTAGG	960
1090	TGCTTTAGTGGCACGGCAACACATTTGAAATGAGTACTACTGGCTGGGTGGAAAGATTTAGG	1149
961	CTTTCAGAACCTCATTTTACCCCTTGTATGGTAATAGTCTTCTACAAGATCATCAGCAATGAA	1020
1150	CTTTCAGAACCTCATTTTACCCCTTGTATGGTAATAGTCTTCTACAAGATCATCAGCAATGAA	1209
1021	GAAGGTTACAGACACATTTGCTATTTCCAAATAGATATAAAGACTGCAATTTATTACA	1080
1210	GAAGGTTACAGACACATTTGCTATTTCCAAATAGATATAAAGACTGCAATTTATTACA	1269
1081	AAAGGCACCTGGGAAGTCACTCGGAGTAGAGCTCTAAACAGTGATTAATCTATCTACATTT	1140
1270	AAAGGCACCTGGGAAGTCACTCGGAGTAGAGCTCTAAACAGTGATTAATCTATCTACATTT	1329
1141	AGTAATGAATATAAGGAATGCCAGGAGGAAGGAATCTTTATATAAATCCAACTTATTGAC	1200
1330	AGTAATGAATATAAGGAATGCCAGGAGGAAGGAATCTTTATATAAATCCAACTTATTGAC	1389
1201	TATACAAAGTGACATGCTCAGTTGTGAGCTGGAATCCGGAAGAGGTGTCACTACTTCT	1260
1390	TATACAAAGTGACATGCTCAGTTGTGAGCTGGAATCCGGAAGAGGTGTCACTACTTCT	1449
1261	GTGTCAATTCAGTAAGAGGGCAAGTATTATCAGCTGAGATGTTCCGGTCTCTGGTCTGCC	1320
1450	GTGTCAATTCAGTAAGAGGGCAAGTATTATCAGCTGAGATGTTCCGGTCTCTGGTCTGCC	1509
1321	CTCTATACTCTPACACAGCAGCGGTGAAATGATAAAGGGCTGAGAGTCTCTGGAGACAATTC	1380
1510	CTCTATACTCTPACACAGCAGCGGTGAAATGATAAAGGGCTGAGAGTCTCTGGAGACAATTC	1569
1381	GCTTTGGATAAAATGTCTGCAGAATGTCCAGATGCCCTCCAAAAAATGGGACTTCATTTATT	1440
1570	GCTTTGGATAAAATGTCTGCAGAATGTCCAGATGCCCTCCAAAAAATGGGACTTCATTTATT	1629

Qy	1441	TTGAATGAACAACAAATTTTGGTATCAGATGATCTTGGCTCCTCATTTTGTGATAAATCCAAG	1500
Db	1630	TTGAATGAACAACAAATTTTGGTATCAGATGATCTTGGCTCCTCATTTTGTGATAAATCCAAG	1689
Qy	1501	AAATATCCTCTACTATTATAGATGTGTATGCAGGCCCATGTAGTCAAAAAGCAGACACGTGTC	1560
Db	1690	AAATATCCTCTACTATTATAGATGTGTATGCAGGCCCATGTAGTCAAAAAGCAGACACGTGTC	1749
Qy	1561	TTCAGACTGAACCTGGGCCACTTACCTTTGCAAGCACAGAAAAACATTTATAGTAGCTAGCTTT	1620
Db	1750	TTCAGACTGAACCTGGGCCACTTACCTTTGCAAGCACAGAAAAACATTTATAGTAGCTAGCTTT	1809
Qy	1621	GATGGCAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGTGCAATCAACAGAAAGACTG	1680
Db	1810	GATGGCAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGTGCAATCAACAGAAAGACTG	1869
Qy	1681	GGAAACATTTCAAGTTGNAGATCAAAATGGAAGCAGCCAGACAAATTTTCAAAAATGGGATTT	1740
Db	1870	GGAAACATTTGAAGTTGAAGATCAAAATGGAAGCAGCCAGACAAATTTTCAAAAATGGGATTT	1929
Qy	1741	GTGGACAAACAAACGAAATTTGGCAATTTGGGGCTGGTCAATGCGAGGGTACGTAAACCTCAATG	1800
Db	1930	GTGGACAAACAAACGAAATTTGGGGCTGGTCAATGCGAGGGTACGTAAACCTCAATG	1989
Qy	1801	GTCTCTGGGATCGGGAAAGTGGCGTTCCTCAAGTGTGGAATAGCCGTGGCGCCTGTATCCCGG	1860
Db	1990	GTCTCTGGGATCGGGAAAGTGGCGTTCCTCAAGTGTGGAATAGCCGTGGCGCCTGTATCCCGG	2049
Qy	1861	TGGGAGTACTATGACTCAGTGTACACAGAACGTTACATGGGTCTCCCAACTCCAGAGAC	1920
Db	2050	TGGGAGTACTATGACTCAGTGTACACAGAACGTTACATGGGTCTCCCAACTCCAGAGAC	2109
Qy	1921	AACCTTGACCATTTACAGAAATTCACACAGTCATGAGCAGAGCTGAAAAATTTTAAACAAGTT	1980
Db	2110	AACCTTGACCATTTACAGAAATTCACACAGTCATGAGCAGAGCTGAAAAATTTTAAACAAGTT	2169
Qy	1981	GAGTACCTCTCTATTTCATGGAACAGCAGATGATAACGTTCACTTTACAGAGTCAGCTCAG	2040
Db	2170	GAGTACCTCTCTATTTCATGGAACAGCAGATGATAACGTTCACTTTACAGAGTCAGCTCAG	2229
Qy	2041	ATCTCCAAAGCCCTGGTTCGATGTTGGAGTGGAATTTCCAGGCAATGTGGTATCTGATGAA	2100
Db	2230	ATCTCCAAAGCCCTGGTTCGATGTTGGAGTGGAATTTCCAGGCAATGTGGTATCTGATGAA	2289
Qy	2101	GACCATGGATAGCTAGCAGCAGACACCAACATATATATACCCACATGAGGCACTTC	2160
Db	2290	GACCATGGATAGCTAGCAGCAGACACCAACATATATATACCCACATGAGGCACTTC	2349
Qy	2161	ATAAAACAAATGTTTCTCTTTACCT	2184
Db	2350	ATAAAACAAATGTTCTCTTTACCT	2373
RESULT 3			
AAL44351			
ID	AAL44351 standard; DNA; 3407 BP.		
XX	AAL44351;		
AC			
XX			
DT	24-OCT-2002 (first entry)		
XX	Human dipeptidyl peptidase IV (DPP IV) coding sequence.		
DE			
XX	Human; angiodemic condition; angiotensin converting enzyme; ACE; gene;		
KW	vasopeptidase inhibitor; dipeptidyl peptidase IV; aminopeptidase P; db;		
KW	DPP IV; aminopeptidase P; APP; hypertension; diabetes; cardiac disease;		
KW	renal disease; enzyme.		
XX			
OS	Homo sapiens.		
XX			
Key	Location/Qualifiers		
FF	76..2376		
FT	CDS		





QY 1501 AATATCTCTACTATTAGATGTGTATGTCAGGCCCATGTAGTCAAAAAGCAGACACTGTC 1560  
DB |||||||  
QY 1690 AATATCTCTACTATTAGATGTGTATGTCAGGCCCATGTAGTCAAAAAGCAGACACTGTC 1749  
DB |||||||  
QY 1561 TTCACTGAACTGGGCCACTTACCTTGCAGCAGACAGAAAACATTTATAGTACTAGCTTTT 1620  
DB |||||||  
QY 1750 TTCAGACTGAATGGGCCACTTACCTTGCAGCAGACAGAAAACATTTATAGTACTAGCTTTT 1809  
DB |||||||  
QY 1621 GATGCAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGTCATCAACAGAGACTG 1680  
DB |||||||  
QY 1810 GATGCAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGTCATCAACAGAGACTG 1869  
DB |||||||  
QY 1681 GGAACATTTGAAGTTGAAGATCAAAATTTGAAGCAGCAGACAAATTTCAAAATGGATTT 1740  
DB |||||||  
QY 1870 GGAACATTTGAAGTTGAAGATCAAAATTTGAAGCAGCAGACAAATTTCAAAATGGATTT 1929  
DB |||||||  
QY 1741 GTGGACAACAAACGAATTTGAGGCTGGTCTATATGGAGGGTACGTAACTCAATG 1800  
DB |||||||  
QY 1930 GTGGAGTACTATGACTCAGTGTACACAGAACTTACATGGGCTGCTCCAACTCCAGAGAC 2109  
DB |||||||  
QY 1921 AACCTTGACCAATACAGAAATTTCAACAGTCAAGCAGAGCTGGAATTTTAAACAAGTT 1980  
DB |||||||  
QY 2110 AACCTTGACCAATACAGAAATTTCAACAGTCAAGCAGAGCTGGAATTTTAAACAAGTT 2169  
DB |||||||  
QY 1981 GAGTACCTCTTATTCATGGAAACAGCAGATGATAACGTTTCACTTTTCAGCAGTCACTGAC 2040  
DB |||||||  
QY 2170 GAGTACCTCTTATTCATGGAAACAGCAGATGATAACGTTTCACTTTTCAGCAGTCACTGAC 2229  
DB |||||||  
QY 2041 ATCTCCAAAGCCCTGGTGCATCTGGAGTGGATTTCCAGGCCAATGGGTATCTGATGAA 2100  
DB |||||||  
QY 2230 ATCTCCAAAGCCCTGGTGCATCTGGAGTGGATTTCCAGGCCAATGGGTATCTGATGAA 2289  
DB |||||||  
QY 2101 GACCATGGAATAGTAGCAGCAGCAGACACCAACATATATATACCCACATGAGCCACTTC 2160  
DB |||||||  
QY 2290 GACCATGGAATAGTAGCAGCAGCAGACACCAACATATATATATACCCACATGAGCCACTTC 2349  
DB |||||||  
QY 2161 ATAAACAATGTTCTCTTTACT 2184  
DB |||||||  
QY 2350 ATAAACAATGTTCTCTTTACT 2373  
DB |||||||

RESULT 4  
ID ABZ22557  
XX standard; cDNA; 3407 BP.  
AC ABZ22557;  
XX  
DT 27-MAR-2003 (first entry)  
DE Human liver dipeptidyl peptidase IV (DPP4) encoding cDNA SEQ ID NO:6.  
XX  
KW Tissue-specific endothelial membrane protein; anticonvulsant; neuroleptic; virucide;  
KW fungicide; antiparasitic; anticonvulsant; neuroleptic; cytostatic;  
KW antiparkinsonian; nootropic; neuroprotective; antiasthmatic; infection;  
KW antidiabetic; hypotensive; nephrotropic; antiarthritic; antiinflammatory;  
KW gene therapy; epilepsy; schizophrenia; cancer; Parkinson's disease;  
KW Alzheimer's disease; asthma; diabetes; hypertension; arthritis;  
KW polycystic kidney disease; tissue-specific luminally expressed protein;  
KW inflammatory bowel disease; enzyme; gene; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT Location/Qualifiers  
CDS 76..2376  
/\*tag= a

FT  
XX  
PN  
XX  
PD  
XX  
PF  
XX  
XX  
07-JUN-2002; 2002WO-US018185.  
08-JUN-2001; 2001US-0297021P.  
PR 12-JUL-2001; 2001US-0305117P.  
XX  
PA (TARG-) TARGET PROTEIN TECHNOLOGIES INC.  
XX  
PI Roben PW, Stevens AC;  
DR WPI; 2003-167367/16.  
XX P-PSDB; ABP56700.  
DR  
DR  
PT Delivering a therapeutic agent (e.g. immunosuppressant) to a specific  
PT tissue, comprises administering an amount of a therapeutic complex  
PT comprising a therapeutic group, a ligand that binds to a tissue-specific  
PT protein, and a linker.  
XX  
PS Claim 24; Page 102-105; 123pp; English.  
XX  
CC The present invention describes a method (M1) for delivering a  
CC therapeutic agent to a specific tissue, comprising administering an  
CC amount of a therapeutic complex comprising a ligand which binds to a  
CC tissue-specific luminally expressed protein, a therapeutic group, and a  
CC linker which links the group to the ligand. Tissue-specific endothelial  
CC membrane proteins from the present invention can have antibacterial,  
CC virucide, fungicide, antiparasitic, anticonvulsant, neuroleptic,  
CC cytostatic, antiparkinsonian, nootropic, neuroprotective, antiasthmatic,  
CC antidiabetic, hypotensive, nephrotropic, antiarthritic and  
CC antiinflammatory activities, and can be used in gene therapy. (M1) is  
CC useful in targeting pharmaceuticals or other therapeutics to specific  
CC tissues using tissue-specific endothelial membrane proteins. A  
CC therapeutic complex may be used to treat or diagnose any disease for  
CC which a tissue- or organ-specific treatment would be efficacious, such as  
CC in cases of infections (e.g. bacterial, viral, fungal and parasitic),  
CC epilepsy, schizophrenia, cancer, Parkinson's disease, Alzheimer's  
CC disease, asthma, diabetes, hypertension, polycystic kidney disease,  
CC arthritis, and inflammatory bowel disease. The present sequence encodes a  
CC human liver dipeptidyl peptidase IV (DPP4), which is used in an example  
CC from the present invention  
XX  
SQ Sequence 3407 BP; 1077 A; 678 C; 704 G; 948 T; 0 U; 0 Other;  
Query Match 100.0%; Score 2184; DB 8; Length 3407;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AGTCGCAAACTTACACTCTAACTGATTACTTTAAAAAATCTTTATAGACTGAAGTTATAC 60  
DB |||||||  
QY 190 AGTCGCAAACTTACACTCTAACTGATTACTTTAAAAAATCTTTATAGACTGAAGTTATAC 249  
DB |||||||  
QY 61 TCCTTAAGATGGATTTTCAGATCATGAATATCTCTACAAACAAGAAAATATCTTGGTA 120  
DB |||||||  
QY 250 TCCTTAAGATGGATTTTCAGATCATGAATATCTCTACAAACAAGAAAATATCTTGGTA 309  
DB |||||||  
QY 121 TTCATGCTGAATATGGAAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 180  
DB |||||||  
QY 310 TTCATGCTGAATATGGAAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 369  
DB |||||||  
QY 181 GGACATCTTATCAATGATTATTCAATATCTCTGATGGGAGTTTATCTCTTAGAATAC 240  
DB |||||||  
QY 370 GGACATCTTATCAATGATTATTCAATATCTCTGATGGGAGTTTATCTCTTAGAATAC 429  
DB |||||||  
QY 241 AACTAGCTGAAGCAATGGAGGCAATCTTACACAGCTTCATATGACATTTATGATTTAAAT 300  
DB |||||||  
QY 430 AACTAGCTGAAGCAATGGAGGCAATCTTACACAGCTTCATATGACATTTATGATTTAAAT 489  
DB |||||||  
QY 301 AAAAGGCAGCTGATTACAGAGAGAGGATTCCAAAACAACACACAGTGGGTACATGGTCA 360  
DB |||||||



PD 31-JUL-2003.  
XX 17-JAN-2003; 2003WO-US001981.  
XX 18-JAN-2002; 2002US-0350061P.  
PR (BRIM ) BRISTOL-MYERS SQUIBB CO.  
PA  
PI Huang F, Fairchild CR, Lee FY, Shaw P;  
XX WPI; 2003-636735/60.  
DR P-PSDB; ADD14045.  
DR  
XX New polynucleotides and polypeptides for predicting the activity of  
PT compounds that interact with protein tyrosine kinases and/or protein  
PT tyrosine kinase pathways.  
XX  
PS Claim 2; SEQ ID NO 34; 139pp; English.  
XX  
CC The present invention describes a predictor set comprising a plurality of  
CC polynucleotides or polypeptides whose expression pattern is predictive of  
CC the response of cells to treatment with a compound that modulates protein  
CC tyrosine kinase activity or members of the protein tyrosine kinase  
CC pathway. Also described: (1) predicting whether a compound is capable of  
CC modulating the activity of cells, comprising obtaining a sample of cells,  
CC determining whether the cells express a plurality of markers, and  
CC correlating the expression of the markers to the compound's ability to  
CC modulate the activity of the cells; (2) a plurality of cell lines for  
CC identifying polynucleotides and polypeptides whose expression levels  
CC correlate with compound sensitivity or resistance of cells associated  
CC with a disease state; and (3) identifying polynucleotides and  
CC polypeptides that predict compound sensitivity or resistance of cells  
CC associated with a disease state, comprising subjecting the plurality of  
CC cell lines to one or more compounds, analysing the expression pattern of  
CC a microarray of polynucleotides or polypeptides, and selecting  
CC polynucleotides or polypeptides that predict the sensitivity or  
CC resistance of cells associated with a disease state by using the  
CC expression pattern of the microarray. The polynucleotides and  
CC polypeptides have cytoskeletal activities, and can be used in gene therapy.  
CC The polynucleotides and polypeptides are useful in predicting the  
CC activity of compounds that interact with protein tyrosine kinases and/or  
CC protein tyrosine kinase pathways. These may be used in determining drug  
CC sensitivity in patients to allow the development of individualized  
CC genetic profiles which aid in treating diseases and disorders (e.g.  
CC cancer) based on patient response at a molecular level. The present  
CC sequence is used in the exemplification of the present invention.  
XX  
SQ Sequence 3407 BP; 1077 A; 678 C; 704 G; 948 T; 0 U; 0 Other;

Query Match 100.0%; Score 2184; DB 10; Length 3407;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTCGCAAACTTACACTCTAACTGATTAATTAATAATACTATTAGACTGAAGTTATAC 60  
DB 190 AGTCGCAAACTTACACTCTAACTGATTAATTAATAATACTATTAGACTGAAGTTATAC 249  
QY 61 TCCTTAAGATGATTTAGATCATGAATATCTCTACAAACAAGAAAATAATATCTTGGA 120  
DB 250 TCCTTAAGATGATTTAGATCATGAATATCTCTACAAACAAGAAAATAATATCTTGGA 309  
QY 121 TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGATTT 180  
DB 310 TTCAATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGATTT 369  
QY 181 GGACATTTCTATCAATGATTTTCAATATCTCTGATGGCAGTTTATTTCTTTAGAATAC 240  
DB 370 GGACATTTCTATCAATGATTTTCAATATCTCTGATGGCAGTTTATTTCTTTAGATAC 429  
QY 241 AACTACGTTGAAGCAATGGAGGCAATCTCTACACAGCTTCTATATGACATTTATGATTTAAAT 300  
DB 430 AACTACGTTGAAGCAATGGAGGCAATCTCTACACAGCTTCTATATGACATTTATGATTTAAAT 489

QY 301 AAAAGGAGCTGATTTACAGAAAGAGAGGATTTCCAAACAACACACAGTGGGTCACTGGTCA 360  
DB 490 AAAAGGAGCTGATTTACAGAAAGAGAGGATTTCCAAACAACACACAGTGGGTCACTGGTCA 549  
QY 361 CCAGTGGGTCAATAATTCGGCATATGTTTGGAAACAATGACATTTATGTTTAAAAATGAACCA 420  
DB 550 CCAGTGGGTCAATAATTCGGCATATGTTTGGAAACAATGACATTTATGTTTAAAAATGAACCA 609  
QY 421 AATTTACCAAGTTACAGAAATCACATGGACGGGGAAGAAGATATATATATATATGGAATA 480  
DB 610 AATTTACCAAGTTACAGAAATCACATGGACGGGGAAGAAGATATATATATATATGGAATA 669  
QY 481 ACTGACTGGGTTTATGAAGAGGAACTCTTCAGTGCCTACTCTGCTCTCTGCTGCTGCTTCCA 540  
DB 670 ACTGACTGGGTTTATGAAGAGGAACTCTTCAGTGCCTACTCTGCTCTCTGCTGCTGCTTCCA 729  
QY 541 AACGGCACTTTTATAGCATATGCCCAATTTAAACGACACAGAAAGTCCCCTTTATGGAATAC 600  
DB 730 AACGGCACTTTTATAGCATATGCCCAATTTAAACGACACAGAAAGTCCCCTTTATGGAATAC 789  
QY 601 TCCTTCTACTCTGATGAGTCACTGAGTACCCAAAGACTGTACGGGTTCATATCCAAAG 660  
DB 790 TCCTTCTACTCTGATGAGTCACTGAGTACCCAAAGACTGTACGGGTTCATATCCAAAG 849  
QY 661 GCAGGAGCTGTGAATCCCAACTGTAAAGTTCTTTTGTAAATAACAGACTCTCTCAGCTCA 720  
DB 850 GCAGGAGCTGTGAATCCCAACTGTAAAGTTCTTTTGTAAATAACAGACTCTCTCAGCTCA 909  
QY 721 GTCCCAATGCAACTTCCATACAAATCACTGCTCTCTGCTTCTATGTTGATAGGGATCAC 780  
DB 910 GTCCCAATGCAACTTCCATACAAATCACTGCTCTCTGCTTCTATGTTGATAGGGATCAC 969  
QY 781 TACTTGTGTGATGACATGAGGCAACACAGAAAGAAATTTCTTTCAGTGGTTCAGGAG 840  
DB 970 TACTTGTGTGATGACATGAGGCAACACAGAAAGAAATTTCTTTCAGTGGTTCAGGAG 1029  
QY 841 ATTCAAGAACTATCGGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900  
DB 1030 ATTCAAGAACTATCGGTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1089  
QY 901 TGCTTAGTGGCAACGCAACACATTTGAAATGAGTACTACTGCTGGTGGTGGTGGTGGTGGT 960  
DB 1090 TGCTTAGTGGCAACGCAACACATTTGAAATGAGTACTACTGCTGGTGGTGGTGGTGGTGGT 1149  
QY 961 CTTTCAGAACCTCATTTTACCTTCATGCTGATGATGATGATGATGATGATGATGATGATGATG 1020  
DB 1150 CTTTCAGAACCTCATTTTACCTTCATGCTGATGATGATGATGATGATGATGATGATGATG 1209  
QY 1021 GAAAGTTACAGACACATTTGCTATTTCCAAATAGATAAAAAGACTGCACATTTATTACA 1080  
DB 1210 GAAAGTTACAGACACATTTGCTATTTCCAAATAGATAAAAAGACTGCACATTTATTACA 1269  
QY 1081 AAAGCACCTCGGGAAGTCAATCGGGAAGAAGCTCTTAACACAGTGAATATCTATATACATTT 1140  
DB 1270 AAAGCACCTCGGGAAGTCAATCGGGAAGAAGCTCTTAACACAGTGAATATCTATATACATTT 1329  
QY 1141 AGTAATGAATATAAAGGAATGCCAGGAGGAAGATCTTTATAAATCCAACTTTATGAC 1200  
DB 1330 AGTAATGAATATAAAGGAATGCCAGGAGGAAGATCTTTATAAATCCAACTTTATGAC 1389  
QY 1201 TATACAAAGTGACATGCTCAGTTGTGAGCTGGAATCCGGAAGGCTGTCAGTACTATTCT 1260  
DB 1390 TATACAAAGTGACATGCTCAGTTGTGAGCTGGAATCCGGAAGGCTGTCAGTACTATTCT 1449  
QY 1261 GTGTCATTTCAAGTAAAGGCGGAAGTATTTATCAGCTGAGATGTTCCGGTCTCTGGTCTGCC 1320  
DB 1450 GTGTCATTTCAAGTAAAGGCGGAAGTATTTATCAGCTGAGATGTTCCGGTCTCTGGTCTGCC 1509  
QY 1321 CTCTATCTCTACACAGCAGGTGAATGATATAAAGGCTGAGAGTCTCTGGAAGACATTTCA 1380  
DB 1510 CTCTATCTCTACACAGCAGGTGAATGATATAAAGGCTGAGAGTCTCTGGAAGACATTTCA 1569  
QY 1381 GCTTTGGGATAAAATGCTGCAGAAATGTCAGATGCCCTCCAAAAAACTGGACTTCTATTATT 1440



QY	121	TTCAATGCTGAATATGGAAACAGCTCAGTTTTCTTGGAGAACAGTACATTTGATGAGTTT	180
DB	310	TTCAATGCTGAATATGGAAACAGCTCAGTTTTCTTGGAGAACAGTACATTTGATGAGTTT	369
QY	181	GGACATTCCTATCAATGATTTTCAATATCTCTGATGGCGAGTTTATCTCTTAGAATAC	240
DB	370	GGACATTCCTATCAATGATTTTCAATATCTCTGATGGCGAGTTTATCTCTTAGAATAC	429
QY	241	AACTACGTGAAGCAATGAGGCATTTCTTACACAGCTTCATATGACATTTATGATTTAAAT	300
DB	430	AACTACGTGAAGCAATGAGGCATTTCTTACACAGCTTCATATGACATTTATGATTTAAAT	489
QY	301	AAAAAGCGACGTGATTTACAGAAAGAGAGGATTCNAAACAACAACAACAGTGGGTACATGTCTCA	360
DB	490	AAAAAGCGACGTGATTTACAGAAAGAGAGGATTCNAAACAACAACAACAGTGGGTACATGTCTCA	549
QY	361	CCAGTGGGTCAATAATTTGGCATATGTTTGGAAACAATGACATTTATGTTTAAATTTGAACCA	420
DB	550	CCAGTGGGTCAATAATTTGGCATATGTTTGGAAACAATGACATTTATGTTTAAATTTGAACCA	609
QY	421	AATTTACCAAGTTACAGAAATCACATGGACGGGGAAGAAGATATAATATAATAGGAATA	480
DB	610	AATTTACCAAGTTACAGAAATCACATGGACGGGGAAGAAGATATAATATAATAGGAATA	669
QY	481	ACTGACTGGGTTTATGAAGAGGAAGTCTTTAGTGCCCTACTCTGCTCTGTGGTGGTCTCCA	540
DB	670	ACTGACTGGGTTTATGAAGAGGAAGTCTTTAGTGCCCTACTCTGCTCTGTGGTGGTCTCCA	729
QY	541	AACGCGCATTTTATAGCATATGCCCAATTTAAACGACACAGAAAGTCCACATTTATTTGAATAC	600
DB	730	AACGCGCATTTTATAGCATATGCCCAATTTAAACGACACAGAAAGTCCACATTTATTTGAATAC	789
QY	601	TCCTTTCTACTCTGATGAGTCACTGCAGTACCCAAAGACTGTATCGGGTTCCATATCCAAAG	660
DB	790	TCCTTTCTACTCTGATGAGTCACTGCAGTACCCAAAGACTGTATCGGGTTCCATATCCAAAG	849
QY	661	GCAGGAGCTGTAATCCAACTGTATAAGTTCTTTGTTGTTAAATAACAGACTCTCTCAGCTCA	720
DB	850	GCAGGAGCTGTAATCCAACTGTATAAGTTCTTTGTTGTTAAATAACAGACTCTCTCAGCTCA	909
QY	721	GTCAACCAATGCAACTTCCATACAAATCACTGCTCTCTCTATGTTGATAGGGGATCCAC	780
DB	910	GTCAACCAATGCAACTTCCATACAAATCACTGCTCTCTCTCTATGTTGATAGGGGATCCAC	969
QY	781	TACTTGTGTGATGTGACATGGGCAACAACAAGAAAGAAATTTCTTTGCAGTGGCTCAGGAGG	840
DB	970	TACTTGTGTGATGTGACATGGGCAACAACAAGAAAGAAATTTCTTTGCAGTGGCTCAGGAGG	1029
QY	841	ATTACAGAACTATTGGTTCATGGATATTTGTGACTATGATGNAATCCAGTGGNAGATGGAAC	900
DB	1030	ATTACAGAACTATTGGTTCATGGATATTTGTGACTATGATGNAATCCAGTGGNAGATGGAAC	1089
QY	901	TGCTTTAGTGGCACGCGCAACAATTTGAAATGAGTACTACTGGCTGGGTTTGGAAAGATTTAGG	960
DB	1090	TGCTTTAGTGGCACGCGCAACAATTTGAAATGAGTACTACTGGCTGGGTTTGGAAAGATTTAGG	1149
QY	961	CCTTTCAGAACTCATTTTACCCCTTGATGGTAAATAGCTTCTACAGAGATCATCAGCAATGAA	1020
DB	1150	CCTTTCAGAACTCATTTTACCCCTTGATGGTAAATAGCTTCTACAGAGATCATCAGCAATGAA	1209
QY	1021	GAAGGTTACAGACACATTTGCTATTTCCNAATAGATATAAAGACTGCAATTTATTTACA	1080
DB	1210	GAAGGTTACAGACACATTTGCTATTTCCAAATAGATATAAAGACTGCAATTTATTTACA	1269
QY	1081	AAAGGCACTCTGGGAAGTCACTCGGGATAGAAGCTCTAAACAGTGATTATCTACTACATTT	1140
DB	1270	AAAGGCACTCTGGGAAGTCACTCGGGATAGAAGCTCTAAACAGTGATTATCTACTACATTT	1329
QY	1141	AGTAATGNAATATAAAGNAATGCCAGGAGGAAGGAATCTTTTATAAAATCCAACTTATTGAC	1200
DB	1330	AGTAATGNAATATAAAGNAATGCCAGGAGGAAGGAATCTTTTATAAAATCCAACTTATTGAC	1389

[illegible]

RESULT 7  
ADO19399  
ID ADO19399 standard; cDNA; 3407 BP.  
XX



AC AD019399;  
 DT 12-AUG-2004 (first entry)  
 DE Human PRO polynucleotide #165.  
 XX  
 KW Human; PRO; gene; ss; immune related disorder;  
 KW systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis;  
 KW juvenile chronic arthritis; systemic sclerosis; Sjogren's syndrome;  
 KW vasculitis; sarcoidosis; autoimmune haemolytic anaemia;  
 KW autoimmune thrombocytopenia; thyroiditis; diabetes mellitus;  
 KW renal disease; demyelinating disease; central nervous system;  
 KW peripheral nervous system; demyelinating polyneuropathy;  
 KW Guillain-Barre syndrome;  
 KW chronic inflammatory demyelinating polyneuropathy.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO2004043361-A2.  
 XX  
 XX 27-MAY-2004.  
 XX  
 XX 06-NOV-2003; 2003WO-US035268.  
 XX  
 XX 08-NOV-2002; 2002US-0425235P.  
 XX  
 XX (GETH ) GENENTECH INC.  
 XX  
 XX Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;  
 XX Wood WI, Wu ID;  
 XX  
 XX WPI; 2004-420067/39.  
 XX P-PSDB; AD019400.  
 XX  
 XX Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for  
 XX treating an immune related disorder such as systemic lupus erythematosus,  
 XX rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or  
 XX spondyloarthritis.  
 XX  
 XX Claim 1; SEQ ID NO 342; 1731pp; English.  
 XX  
 XX The invention relates to human PRO polypeptides and the polynucleotides  
 XX encoding them. The polypeptides and polynucleotides are useful for  
 XX treating and diagnosing immune related disorders in mammals. The immune  
 XX related disorders include systemic lupus erythematosus, rheumatoid  
 XX arthritis, osteoarthritis, juvenile chronic arthritis, systemic  
 XX sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune  
 XX haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes  
 XX mellitus, immune-mediated renal disease, demyelinating diseases of the  
 XX central or peripheral nervous system, demyelinating polyneuropathy,  
 XX Guillain-Barre syndrome and chronic inflammatory demyelinating  
 XX polyneuropathy. This sequence represents a human PRO polynucleotide of  
 XX the invention.  
 XX  
 XX Sequence 3407 BP; 1077 A; 678 C; 704 G; 948 T; 0 U; 0 Other;  
 XX  
 XX Query Match 100.0%; Score 2184; DB 12; Length 3407;  
 XX Best Local Similarity 100.0%; Pred. No. 0;  
 XX Matches 2184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 Qy 1 AGTCGCAAACTTACACTCTAACTGATTACTTAAAAAATACTTATAGACTGAAGTTATAC 60  
 Db 190 AGTCGCAAACTTACACTCTAACTGATTACTTAAAAAATACTTATAGACTGAAGTTATAC 249  
 Qy 61 TCCTTAAGATGGATTTCAGATCATGATATCTCTACAAAACAGAAAATAATATCTTGGTA 120  
 Db 250 TCCTTAAGATGGATTTCAGATCATGATATCTCTACAAAACAGAAAATAATATCTTGGTA 309  
 Qy 121 TTCATCTGTAATGGAACAGCTCAGTTTCTTTGGAGAACACTACATTTGATGAGTTT 180  
 Db 310 TTCATCTGTAATGGAACAGCTCAGTTTCTTTGGAGAACACTACATTTGATGAGTTT 369  
 Qy 181 GGACATTTCTATCAATGATTAATTAATCTCTCTGATGGCAGTTTATTTCTTCTAGATAC 240

Db 370 GGACATTTCTATCAATGATTAATTAATCTCTGATGGCAGTTTATTTCTTCTAGATAC 429  
 Qy 241 AACTACGTGAAGCAATGGAGCATTCTCTACACAGCTTCATATGATTAATTAATTAAT 300  
 Db 430 AACTACGTGAAGCAATGGAGCATTCTCTACACAGCTTCATATGATTAATTAATTAAT 489  
 Qy 301 AAAAGGCACTGATTATACAGAGAGAGGATTCCTCAACACACACAGTGGGTCACTGGTCA 360  
 Db 490 AAAAGGCACTGATTATACAGAGAGAGGATTCCTCAACACACACAGTGGGTCACTGGTCA 549  
 Qy 361 CCAGTGGGTCAATAATTTGGCATATGTTTGGAAACAATGACATTTATTTAAATTTGAACA 420  
 Db 550 CCAGTGGGTCAATAATTTGGCATATGTTTGGAAACAATGACATTTATTTAAATTTGAACA 609  
 Qy 421 AATTTACCAAGTTACAGAAATCAGATGACGCGGGAAGAGATATAATATAATTAATGGAATA 480  
 Db 610 AATTTACCAAGTTACAGAAATCAGATGACGCGGGAAGAGATATAATATAATTAATGGAATA 669  
 Qy 481 ACTGACTGGGTTTATGAAGAGAGAGTCTTTCAGTGCCTACTCTGCTCTGCTGGTGGTCTCCA 540  
 Db 670 ACTGACTGGGTTTATGAAGAGAGAGTCTTTCAGTGCCTACTCTGCTCTGCTGGTGGTCTCCA 729  
 Qy 541 AACGGCACTTTTATGATATGCCCCAATTTTAAACACACAGAAAGTCCCACTTATTTGAATAC 600  
 Db 730 AACGGCACTTTTATGATATGCCCCAATTTTAAACACACAGAAAGTCCCACTTATTTGAATAC 789  
 Qy 601 TCCTTCTACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
 Db 790 TCCTTCTACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 849  
 Qy 661 GCAGGAGCTGTGAATCCAACTGTAAAGTCTTCTTGTGTAAATACAGACTCTCTCAGCTCA 720  
 Db 850 GCAGGAGCTGTGAATCCAACTGTAAAGTCTTCTTGTGTAAATACAGACTCTCTCAGCTCA 909  
 Qy 721 GTCAACCAATGCAACTTCCATACAAATCACTGCTCTCTCTATGTTGATAGGAGATCAC 780  
 Db 910 GTCAACCAATGCAACTTCCATACAAATCACTGCTCTCTCTATGTTGATAGGAGATCAC 969  
 Qy 781 TACTTGTGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840  
 Db 970 TACTTGTGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1029  
 Qy 841 ATTACAGAACTATTCCGTCATGGATATTTGTGACTATGATGAATCCAGTGGAGATGGAAC 900  
 Db 1030 ATTACAGAACTATTCCGTCATGGATATTTGTGACTATGATGAATCCAGTGGAGATGGAAC 1089  
 Qy 901 TGCTTACTGGCAGGCAACACATTTGAATGAGTACTACTGCTGGGTGGAGATTTTAGG 960  
 Db 1090 TGCTTACTGGCAGGCAACACATTTGAATGAGTACTACTGCTGGGTGGAGATTTTAGG 1149  
 Qy 961 CCTTCAGAACTCATTTTACCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
 Db 1150 CCTTCAGAACTCATTTTACCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1209  
 Qy 1021 GAAGTTTACAGACACATTTTCTTCCAAATAGATAAAAAGAGACTGCACATTTATTTACA 1080  
 Db 1210 GAAGTTTACAGACACATTTTCTTCCAAATAGATAAAAAGAGACTGCACATTTATTTACA 1269  
 Qy 1081 AAAGGCACTTGGGAGAGTCACTGGGATAGAGCTTAAACAGTGTATTTACTACTACTACT 1140  
 Db 1270 AAAGGCACTTGGGAGAGTCACTGGGATAGAGCTTAAACAGTGTATTTACTACTACTACT 1329  
 Qy 1141 AGTAATGAATATAAAGGAATGCCAGGAGGAGGATCTTTTATAAATCCCACTTATTTGAC 1200  
 Db 1330 AGTAATGAATATAAAGGAATGCCAGGAGGAGGATCTTTTATAAATCCCACTTATTTGAC 1389  
 Qy 1201 TATACAAAAGTGCATGCTCAGTTGTGAGTGTGAATCCGGAAGAGTGTGAGTACTATTCT 1260  
 Db 1390 TATACAAAAGTGCATGCTCAGTTGTGAGTGTGAATCCGGAAGAGTGTGAGTACTATTCT 1449  
 Qy 1261 GTGTCTTCACTTAAAGAGGCAAGTATTTATCAGCTGAGATGTTTCCGGTCTGGTCTGCC 1320





Db 250 TCCTTAAGATGGAATTCAGATCATGAATATCTCTACAAACAGAAAAATAATATCTTGTA 309  
 QY 121 TTCAATGCTGAATATGGAACAGCTCAGTTTTCTTGGAGAACAGTACATTTTATGATGATTT 180  
 Db 310 TTCAATGCTGAATATGGAACAGCTCAGTTTTCTTGGAGAACAGTACATTTTATGATGATTT 369  
 QY 181 GGAATTTCTATCAATGATTTATTAATATCTCTGATGGCAGTTTATTTCTCTTGAATAC 240  
 Db 370 GGAATTTCTATCAATGATTTATTAATATCTCTGATGGCAGTTTATTTCTCTTGAATAC 429  
 QY 241 AACTACGTGAGCAATGAGGATTCCTACAGCTTCATATGATGATTTATTTATTTAAT 300  
 Db 430 AACTACGTGAGCAATGAGGATTCCTACAGCTTCATATGATGATTTATTTAATTTAAT 489  
 QY 301 AAAAGGAGCTGATTTACAGAGAGAGGATTCCTACAGCTTCATATGATGATTTATTTAAT 360  
 Db 490 AAAAGGAGCTGATTTACAGAGAGAGGATTCCTACAGCTTCATATGATGATTTATTTAAT 549  
 QY 361 CCAGTGGGTCAATAATGGCATAATGTTGGCAATGATGATTTATTTAATTTGAACCA 420  
 Db 550 CCAGTGGGTCAATAATGGCATAATGTTGGCAATGATGATTTATTTAATTTGAACCA 609  
 QY 421 AATTTACCAAGTTACAGATTCATGAGAGGAGGATTCCTACAGCTTCATATGATGATTTAAT 480  
 Db 610 AATTTACCAAGTTACAGATTCATGAGAGGAGGATTCCTACAGCTTCATATGATGATTTAAT 669  
 QY 481 ACTGACTGGGTTTATGAAGAGGAGTCTTTCAGTGGCTACTCTGCTGCTGGTGGTCTCCA 540  
 Db 670 ACTGACTGGGTTTATGAAGAGGAGTCTTTCAGTGGCTACTCTGCTGCTGGTGGTCTCCA 729  
 QY 541 AACGGCACTTTTTAGCATATGCCCAATTTAAACGACACAGAGTCCCACTTTATTTGAATAC 600  
 Db 730 AACGGCACTTTTTAGCATATGCCCAATTTAAACGACACAGAGTCCCACTTTATTTGAATAC 789  
 QY 601 TCCTTCTACTCTGATGAGTCACTGAGTACCCAAAGACTGACGGGTTCCATATCCAAAG 660  
 Db 790 TCCTTCTACTCTGATGAGTCACTGAGTACCCAAAGACTGACGGGTTCCATATCCAAAG 849  
 QY 661 GCAGGAGCTGTGAATCCCAATGTAAGTTCTTTGTTGTAATACAGACTCTCTCAGCTCA 720  
 Db 850 GCAGGAGCTGTGAATCCCAATGTAAGTTCTTTGTTGTAATACAGACTCTCTCAGCTCA 909  
 QY 721 GTCACCAATGCAATTCATACAAATCACTGCTCTGCTCTCTATGTTGATAGGGGATCAC 780  
 Db 910 GTCACCAATGCAATTCATACAAATCACTGCTCTGCTCTCTATGTTGATAGGGGATCAC 969  
 QY 781 TACTTGTGATGATGACATGGGCAACACAGAAAGAAATTTCTTTCAGTGGCTCAGGAGG 840  
 Db 970 TACTTGTGATGATGACATGGGCAACACAGAAAGAAATTTCTTTCAGTGGCTCAGGAGG 1029  
 QY 841 ATTCAGAACTATTCCGTCATGATATTTGTCATGATGATGATGATGATGATGATGATGATG 900  
 Db 1030 ATTCAGAACTATTCCGTCATGATATTTGTCATGATGATGATGATGATGATGATGATGATG 1089  
 QY 901 TGCTTAGTGGAACGGCAACATTTGAATGATGATGATGATGATGATGATGATGATGATGATG 960  
 Db 1090 TGCTTAGTGGAACGGCAACATTTGAATGATGATGATGATGATGATGATGATGATGATGATG 1149  
 QY 961 CCTTCAGAACCTCAATTTTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATG 1020  
 Db 1150 CCTTCAGAACCTCAATTTTACCTTGAATGATGATGATGATGATGATGATGATGATGATGATG 1209  
 QY 1021 GAAAGTTACAGACATTTGCTATTTCCAAATAGTAAAGAACTGACACATTTATTTACA 1080  
 Db 1210 GAAAGTTACAGACATTTGCTATTTCCAAATAGTAAAGAACTGACACATTTATTTACA 1269  
 QY 1081 AAAGGCACTGGGAGTCACTGGGATAGAGCTCTAACCAAGTATGATTTATTTATTTACATTT 1140  
 Db 1270 AAAGGCACTGGGAGTCACTGGGATAGAGCTCTAACCAAGTATGATTTATTTATTTACATTT 1329  
 QY 1141 AGTAATGAATATAAGGAATGCCAGGAGGAAGGAATCTTTTATAAAATCCCACTTTATTGAC 1200

Db 1330 AGTAATGAATATAAGGAATGCCAGGAGGAAGGAATCTTTTATAAAATCCAACTTATTGAC 1389  
 QY 1201 TATACAAAGTGACATGCTCAGTTGTGAGCTGGAATCCGGAAGAGTGTGAGTACTATTCT 1260  
 Db 1390 TATACAAAGTGACATGCTCAGTTGTGAGCTGGAATCCGGAAGAGTGTGAGTACTATTCT 1449  
 QY 1261 GTGTCAATTCAGTAAAGAGGCGAAGTATTTATCAGCTGAGATGTTCCGGTCTGCTGCTGCC 1320  
 Db 1450 GTGTCAATTCAGTAAAGAGGCGAAGTATTTATCAGCTGAGATGTTCCGGTCTGCTGCTGCC 1509  
 QY 1321 CTCTATCTCTACACAGCAGCGTGAATGAATGAAGGCTGAGAGTCTCTGGAAGACAATTTCA 1380  
 Db 1510 CTCTATCTCTACACAGCAGCGTGAATGAATGAAGGCTGAGAGTCTCTGGAAGACAATTTCA 1569  
 QY 1381 GCTTTGGATAAAATGCTGCAGAAATGCTCAGATGCTCCAAAATACTGGAATCTTCAATTAT 1440  
 Db 1570 GCTTTGGATAAAATGCTGCAGAAATGCTCAGATGCTCCAAAATACTGGAATCTTCAATTAT 1629  
 QY 1441 TTGAATGAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGTATAAAATCCAAAG 1500  
 Db 1630 TTGAATGAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGTATAAAATCCAAAG 1689  
 QY 1501 AAATATCTCTACTACTATTTAGATGTTATGTCAGGCCCCATGTCACAAAAGCAGACACTGTC 1560  
 Db 1690 AAATATCTCTACTACTATTTAGATGTTATGTCAGGCCCCATGTCACAAAAGCAGACACTGTC 1749  
 QY 1561 TTCAAGCTGAACTGGGCGACCTTACCTTGCAGACACAGAAAACATTTATAGTAGTACTGCTTT 1620  
 Db 1750 TTCAAGCTGAACTGGGCGACCTTACCTTGCAGACACAGAAAACATTTATAGTAGTACTGCTTT 1809  
 QY 1621 GATGGCAGAGGAAGTGGTTACCAAGGAGATAAGATCATGCAATGCAATCAACAGAGAGACTG 1680  
 Db 1810 GATGGCAGAGGAAGTGGTTACCAAGGAGATAAGATCATGCAATCAACAGAGAGACTG 1869  
 QY 1681 GGAACATTTGAGTTGAAGTCAAAATTTGAGCAGCAGCAGACAAATTTTCAAAAATGGGATTT 1740  
 Db 1870 GGAACATTTGAGTTGAAGTCAAAATTTGAGCAGCAGCAGACAAATTTTCAAAAATGGGATTT 1929  
 QY 1741 GTGAGCAACAAACGAAATTTGCAATTTGGGCTGGTGCATATGAGGAGTACGTAACCTCAATG 1800  
 Db 1930 GTGAGCAACAAACGAAATTTGCAATTTGGGCTGGTGCATATGAGGAGTACGTAACCTCAATG 1989  
 QY 1801 GTCCTGGGATGGGAAGTGGGCTGTTCAAGTGTGGAATAGCCGCTGGGCTGTATCCCGG 1860  
 Db 1990 GTCCTGGGATGGGAAGTGGGCTGTTCAAGTGTGGAATAGCCGCTGGGCTGTATCCCGG 2049  
 QY 1861 TGGGAGTACTATGACTCAGTGTACACAGAACGTTTACATGGGTCTCCCACTCCAGAGAC 1920  
 Db 2050 TGGGAGTACTATGACTCAGTGTACACAGAACGTTTACATGGGTCTCCCACTCCAGAGAC 2109  
 QY 1921 AACCTTGACCAATTTACAGAAATTTCAACAGTATGAGCAGAGCTGAAAATTTTAAACAAGTT 1980  
 Db 2110 AACCTTGACCAATTTACAGAAATTTCAACAGTATGAGCAGAGCTGAAAATTTTAAACAAGTT 2169  
 QY 1981 GAGTACCTCTTATTTATGGAACAGCAGATGATTAACCTTCACTTTCAGCAGTCACTCAG 2040  
 Db 2170 GAGTACCTCTTATTTATGGAACAGCAGATGATTAACCTTCACTTTCAGCAGTCACTCAG 2229  
 QY 2041 ATCTCCAAAGCCCTGGTTCGATGTTGGAGTGGATTTCCAGGCAATGTGGTATCTGATGAA 2100  
 Db 2230 ATCTCCAAAGCCCTGGTTCGATGTTGGAGTGGATTTCCAGGCAATGTGGTATCTGATGAA 2289  
 QY 2101 GACCATCGAATAGCTAGCAGCAGACACCAACATATATATATATATATATATATATATATATAT 2160  
 Db 2290 GACCATCGAATAGCTAGCAGCAGACACCAACATATATATATATATATATATATATATATATAT 2349  
 QY 2161 ATAAAAACAATGTTTCTCTTTACCT 2184  
 Db 2350 ATAAAAACAATGTTTCTCTTTACCT 2373





Query Match 99.9%; Score 2182.4; DB 12; Length 2301; Best Local Similarity 99.9%; Pred. No. 0; Matches 2183; Conservative 0; Mismatches 1; Indels 0; Gaps 0;									
QY	1	AGTCGCAAACTTACACTCTAATGATTAATCTTAAATAATCTTATAGACTGAAGTTATAC	60	QY	1021	GAAGTTACAGACACATTTGCTATTTCCAAATAGATAAAAAAGACTGCACATTTATTACA	1080		
DB	115	AGTCGCAAACTTACACTCTAATGATTAATCTTAAATAATCTTATAGACTGAAGTTATAC	174	DB	1135	GAAGTTACAGACACATTTGCTATTTCCAAATAGATAAAAAAGACTGCACATTTATTACA	1194		
QY	61	TCCTTAAGATGATTTACAGATCATGATATCTCTACAAACAGAAATAATATCTTGGTA	120	QY	1081	AAAGGCACCTGGGAAGTCTCGGATAGAAGCTCTAACACAGTGAATATCTATACTACATT	1140		
DB	175	TCCTTAAGATGATTTACAGATCATGATATCTCTACAAACAGAAATAATATCTTGGTA	234	DB	1195	AAAGGCACCTGGGAAGTCTCGGATAGAAGCTCTAACACAGTGAATATCTATACTACATT	1254		
QY	121	TTCAATGCTGAATATGGAAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT	180	QY	1141	AGTAATGAATATAAAGGAATGCCAGGAGGAAGGAATCTTTATAAAATCCAACTTATTGAC	1200		
DB	235	TTCAATGCTGAATATGGAAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT	294	DB	1255	AGTAATGAATATAAAGGAATGCCAGGAGGAAGGAATCTTTATAAAATCCAACTTATTGAC	1314		
QY	181	GGACATTTCTAATGATTTTCAATATCTCTGATGGGCGAGTTTATCTCTTAGAATAC	240	QY	1201	TATACAAAGTGACATGCCTCAGTTGTGAGCTGAATCCGGAAGAGTGCTAGTACTATTCT	1260		
DB	295	GGACATTTCTAATGATTTTCAATATCTCTGATGGGCGAGTTTATCTCTTAGAATAC	354	DB	1315	TATACAAAGTGACATGCCTCAGTTGTGAGCTGAATCCGGAAGAGTGCTAGTACTATTCT	1374		
QY	241	AACTAGCTGAAGCAATGGAGGCAATCTCTACACAGCTTCATATGACATTTTATGAAT	300	QY	1261	GTGTCAATTCAGTAAAGAGGCGGAAGTATTATCAGCTGAGATGTTCCGGTCTTGCTCGCC	1320		
DB	355	AACTAGCTGAAGCAATGGAGGCAATCTCTACACAGCTTCATATGACATTTTATGAAT	414	DB	1375	GTGTCAATTCAGTAAAGAGGCGGAAGTATTATCAGCTGAGATGTTCCGGTCTTGCTCGCC	1434		
QY	301	AAAGGCAGCTGATTTACAGAGAGAGGATTCCAAACAAACACACAGTGGGTCAATGGTCA	360	QY	1321	CTCTATATCTCTACACAGCAGCGTGAATGATAAAGGGCTGAGAGTCTCTGGAAGACAAATCA	1380		
DB	415	AAAGGCAGCTGATTTACAGAGAGAGGATTCCAAACAAACACACAGTGGGTCAATGGTCA	474	DB	1435	CTCTATATCTCTACACAGCAGCGTGAATGATAAAGGGCTGAGAGTCTCTGGAAGACAAATCA	1494		
QY	361	CCAGTGGGTTCATAAATGGCATAATGTTTGGAAACAATGACATTTATGTTAAATTTGAACCA	420	QY	1381	GCTTTGGATAAAATGCTGCAGAAATGTCAGATGCCCTCCAAAAAACTGGACTTCAATTATT	1440		
DB	475	CCAGTGGGTTCATAAATGGCATAATGTTTGGAAACAATGACATTTATGTTAAATTTGAACCA	534	DB	1495	GCTTTGGATAAAATGCTGCAGAAATGTCAGATGCCCTCCAAAAAACTGGACTTCAATTATT	1554		
QY	421	AATTTACCAAGTTACAGAAATCAGTGCAGCGGGAAGAGATATAATATATATGGAATA	480	QY	1441	TTGAATGAAACAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGTATTAATCCAG	1500		
DB	535	AATTTACCAAGTTACAGAAATCAGTGCAGCGGGAAGAGATATAATATATATGGAATA	594	DB	1555	TTGAATGAAACAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGTATTAATCCAG	1614		
QY	481	ACTGACTGGGTTTATCAAGAGGAAGTCTTCAGTGCGCTACTCTGCTCTGTTGGTGTCTCCA	540	QY	1501	AAATATCTCTACTATTAGATGTGTATGACGCCCATGTAGTCAAAAAAGCAGACACTGTC	1560		
DB	595	ACTGACTGGGTTTATCAAGAGGAAGTCTTCAGTGCGCTACTCTGCTCTGTTGGTGTCTCCA	654	DB	1615	AAATATCTCTACTATTAGATGTGTATGACGCCCATGTAGTCAAAAAAGCAGACACTGTC	1674		
QY	541	AACGGCACTTTTATAGCATATGCCAAATTTAACGACACAGAAAGTCCCACTTATGAATAC	600	QY	1561	TTACAGACTGAACTGGGCCACTTTACCTTGCAAGCACAGAAAAACATTTATAGTAGCTTTT	1620		
DB	655	AACGGCACTTTTATAGCATATGCCAAATTTAACGACACAGAAAGTCCCACTTATGAATAC	714	DB	1675	TTACAGACTGAACTGGGCCACTTTACCTTGCAAGCACAGAAAAACATTTATAGTAGCTTTT	1734		
QY	601	TCCTTCTACTGTGATGAGTCACTGCAGTACCCAAAGACTGTACGGGTTCCATATCCAAAG	660	QY	1621	GATGCGAGGAAAGTGGTTTACCAAGGAGATAAGATCATGATGCAATCAACAGAGACTG	1680		
DB	715	TCCTTCTACTGTGATGAGTCACTGCAGTACCCAAAGACTGTACGGGTTCCATATCCAAAG	774	DB	1735	GATGCGAGGAAAGTGGTTTACCNAAGGAGATAAGATCATGATGCAATCAACAGAGACTG	1794		
QY	661	GCAGGAGCTGTAATCCAACTGTAAAGTTCTTTGTTGTAATAACAGACTCTCTCAGCTCA	720	QY	1681	GGAAACATTTGAAGTTTGAAGATCAAAATTTGAAGCAGCCAGACAAATTTTCAAAAAATGGGAATTT	1740		
DB	775	GCAGGAGCTGTGNAATCCAACTGTAAAGTTCTTTGTTGTAATAACAGACTCTCTCAGCTCA	834	DB	1795	GGAACTTTGAAGTTTGAAGATCAAAATTTGAAGCAGCCAGACAAATTTTCAAAAAATGGGAATTT	1854		
QY	721	GTCAACAAATGCCAACTTCCATACAAATCACTGCTCTGCTTCTATGTTGATAGGGGATCAC	780	QY	1741	GTGGACAAACAAACGAATTTGGGGCTGGTCAATATGGAGGGTACGTAACCTCAATG	1800		
DB	835	GTCAACAAATGCCAACTTCCATACAAATCACTGCTCTGCTTCTATGTTGATAGGGGATCAC	894	DB	1855	GTGGACAAACAAACGAATTTGGGGCTGGTCAATATGGAGGGTACGTAACCTCAATG	1914		
QY	781	TACTTGTGTGATGTGACATGGGCAACACAAAGAAAGAAATTTCTTGTGAGTGGCTCAGAGG	840	QY	1801	GTCTGGGATCGGGAAGTGGCGTGTTCAGTGTGGAATAGCCGTGGCGCTGTATCCCGG	1860		
DB	895	TACTTGTGTGATGTGACATGGGCAACACAAAGAAAGAAATTTCTTGTGAGTGGCTCAGAGG	954	DB	1915	GTCTGGGATCGGGAAGTGGCGTGTTCAGTGTGGAATAGCCGTGGCGCTGTATCCCGG	1974		
QY	841	ATTCAGAACTATTCGGTTCATGATATTTGTGACTATGATGAATCCAGTGGAGATGGAAAC	900	QY	1861	TGGGAGTACTATGACTCAGTGTACACAGAACTTATGAGGTCTCCCACTCCAGAGAC	1920		
DB	955	ATTCAGAACTATTCGGTTCATGATATTTGTGACTATGATGAATCCAGTGGAGATGGAAAC	1014	DB	1975	TGGGAGTACTATGACTCAGTGTACACAGAACTTATGAGGTCTCCCACTCCAGAGAC	2034		
QY	901	TGCTTAGTGGCAGGCAACACATTTGAAATGAGTACTACTGGCTGGGTTGGAAGATTTAGG	960	QY	1921	AACCTTTGACCAATACAGAAATTCACACAGTCTAGCAGAGCTGAAAAATTTTAAACAGTT	1980		
DB	1015	TGCTTAGTGGCAGGCAACACATTTGAAATGAGTACTACTGGCTGGGTTGGAAGATTTAGG	1074	DB	2035	AACCTTTGACCAATACAGAAATTCACACAGTCTAGCAGAGCTGAAAAATTTTAAACAGTT	2094		
QY	961	CCTTCAGAACTCATTTTACCTTTGATGGTAAATAGCTTCTACAAAGATCATCAGCAATGAA	1020	QY	1981	GAGTACCTCTCTTATTCATGGAAACAGCAGATGATAACGTTTCACTTTCCAGAGCTCAG	2040		
DB	1075	CCTTCAGAACTCATTTTACCTTTGATGGTAAATAGCTTCTACAAAGATCATCAGCAATGAA	1134	DB	2095	GAGTACCTCTCTTATTCATGGAAACAGCAGATGATAACGTTTCACTTTCCAGAGCTCAG	2154		

[illegible]





```
QY 241 AACTACGTGAAGCAATGAGGCAATTCCTACACAGCTTCATATGACATTTATGATTTAAAT 300
Db 380 AACTACGTGAAGCAATGAGGCAATTCCTACACAGCTTCATATGACATTTATGATTTAAAT 439
QY 301 AAAAGGCAAGCTGATTTACAGAGAGAGAGATTCCTAACACACACACAGTGGGTCACTGGTCA 360
Db 440 AAAAGGCAAGCTGATTTACAGAGAGAGAGATTCCTAACACACACACAGTGGGTCACTGGTCA 499
QY 361 CCAGTGGGTCAATAATTTGGGCATATGTTTGGAAACAATGACATTTATGTTAAAAATTTGAACCA 420
Db 500 CCAGTGGGTCAATAATTTGGGCATATGTTTGGAAACAATGACATTTATGTTAAAAATTTGAACCA 559
QY 421 AATTATCAAGATTACAGAAATCAATGGAACGGGGAAGAAAGATATAAATATAAATGAATA 480
Db 560 AATTATCAAGATTACAGAAATCAATGGAACGGGGAAGAAAGATATAAATATAAATGAATA 619
QY 481 ACTGACTGGGTTTATGAAGAGAGAGTCTTCAGTGGCTACTCTGCTCTGCTGGTGGTCTCCA 540
Db 620 ACTGACTGGGTTTATGAAGAGAGAGTCTTCAGTGGCTACTCTGCTCTGCTGGTGGTCTCCA 679
QY 541 AACGGCACTTTTATAGCATATGCCCAATTTTAAACGACACAGAAAGTCCCACTTTATTTGAATAC 600
Db 680 AACGGCACTTTTATAGCATATGCCCAATTTTAAACGACACAGAAAGTCCCACTTTATTTGAATAC 739
QY 601 TCCTTCTACTCTGATGAGTCACTGCAATGCCAAGACTGTAACGGGTTCCATATCCAAAG 660
Db 740 TCCTTCTACTCTGATGAGTCACTGCAATGCCAAGACTGTAACGGGTTCCATATCCAAAG 799
QY 661 GCAGGAGTGTGAATCCCACTGTAAAGTCTTTGTTGTTAAATACAGACTCTCTCAGCTCA 720
Db 800 GCAGGAGTGTGAATCCCACTGTAAAGTCTTTGTTGTTAAATACAGACTCTCTCAGCTCA 859
QY 721 GTCCACCAATGCAATCTTCATACAAATCACTGCTCCTGCTCTATGTTGATAGGGGATCAC 780
Db 860 GTCCACCAATGCAATCTTCATACAAATCACTGCTCCTGCTCTATGTTGATAGGGGATCAC 919
QY 781 TACTTGTGTGATGTGACATGGGCAACACAGAAAGAAATTTCTTTGCACTGGCTCAGGAGG 840
Db 920 TACTTGTGTGATGTGACATGGGCAACACAGAAAGAAATTTCTTTGCACTGGCTCAGGAGG 979
QY 841 ATTGAGAACTATTTGGTCAATGATATTTGATGATGATGATGATGATGATGATGATGATGATG 900
Db 980 ATTGAGAACTATTTGGTCAATGATATTTGATGATGATGATGATGATGATGATGATGATGATG 1039
QY 901 TGCTTAGTGGCAGCGCACACATTTGAATGAGTACTCTGCTGGTGGTGGAGATTTTAGG 960
Db 1040 TGCTTAGTGGCAGCGCACACATTTGAATGAGTACTCTGCTGGTGGTGGAGATTTTAGG 1099
QY 961 CCTTCAGAACTCATTTTACCTTTGATGGTAAATAGCTTTCTACAAGATCATCAGCAATGAA 1020
Db 1100 CCTTCAGAACTCATTTTACCTTTGATGGTAAATAGCTTTCTACAAGATCATCAGCAATGAA 1159
QY 1021 GAAGGTTACAGACATTTGCTATTTTCCAAATAGATAAATAAAGACTGCACATTTATTTACA 1080
Db 1160 GAAGGTTACAGACATTTGCTATTTTCCAAATAGATAAATAAAGACTGCACATTTATTTACA 1219
QY 1081 AAAGSCACTGGGAGTCACTGGGATAGAAGCTCTAACCACTGATTTATCTACTACATT 1140
Db 1220 AAAGSCACTGGGAGTCACTGGGATAGAAGCTCTAACCACTGATTTATCTACTACATT 1279
QY 1141 AGTAATGAATATAAGGAATGCCAGGAGGAAGGAATCTTTATAAAATCCAACTTTATTGAC 1200
Db 1280 AGTAATGAATATAAGGAATGCCAGGAGGAAGGAATCTTTATAAAATCCAACTTTATTGAC 1339
QY 1201 TATACAAAGTGACATGCTCAGTTGTGAGTGTGAATCCGGAAAGGTGTCACTATTTCCT 1260
Db 1340 TATACAAAGTGACATGCTCAGTTGTGAGTGTGAATCCGGAAAGGTGTCACTATTTCCT 1399
QY 1261 GTGTCAATTCAGTAAAGGCGCAAGTATTTATCAGCTGAGATGTTCCGGTCCCTGGTCTGCC 1320
Db 1400 GTGTCAATTCAGTAAAGGCGCAAGTATTTATCAGCTGAGATGTTCCGGTCCCTGGTCTGCC 1459
QY 1321 CTCTATACCTACACAGCAGCGGTGAATGATAAAGGGCTGAGAGTCTCGGAAAGACAAATTC 1380
```

RESULT 13  
ADJ74790

ID ADJ74790 standard; DNA; 3445 BP.

XX AC ADJ74790;

XX DT 20-MAY-2004 (first entry)

XX DE Marker gene SEQ ID NO:42.

XX KW bronchial asthma; chronic obstructive pulmonary disease;

KW respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;

KW gene therapy; marker gene; gene; ds.

```
Db 1460 CTCTATATCTTACACAGCAGCGTGAATGATAAAGGGCTGAGAGTCTCTGGAAGACAAATTC 1519
QY 1381 GCTTTGGATATAAATGCTTGCAGAAATGTCAGATGCCCTCCAAAAAATCGCACTTCATTTATT 1440
Db 1520 GCTTTGGATATAAATGCTTGCAGAAATGTCAGATGCCCTCCAAAAAATCGCACTTCATTTATT 1579
QY 1441 TTGAATGAAAAAATAAATTTGGTATCAGATGATCTTGGCTCTCTCAATTTTGTATAAATCCAAAG 1500
Db 1580 TTGAATGAAAAAATAAATTTGGTATCAGATGATCTTGGCTCTCTCAATTTTGTATAAATCCAAAG 1639
QY 1501 AAATATCCTCTACTATTTAGATGTTGTCAGGCCCATGTAGTCAAAAAGCAGACACTGTC 1560
Db 1640 AAATATCCTCTACTATTTAGATGTTGTCAGGCCCATGTAGTCAAAAAGCAGACACTGTC 1699
QY 1561 TTCAAGACTGAACTCGGGCCACTTACCTTGGCAAGCAGAAAAACATTTATAGTAGCTAGCTTT 1620
Db 1700 TTCAAGACTGAACTCGGGCCACTTACCTTGGCAAGCAGAAAAACATTTATAGTAGCTAGCTTT 1759
QY 1621 GATGGCAGAGGAAGTGGTTTACCAAGGAGATTAAGATCATGCAATGCAATCAACAGAAAGCTG 1680
Db 1760 GATGGCAGAGGAAGTGGTTTACCAAGGAGATTAAGATCATGCAATCAACAGAAAGCTG 1819
QY 1681 GGAACATTTGAAGTTGAAGATCAAAATTTGAAGCAGCCAGACAAATTTTCAAAAATGGGATTT 1740
Db 1820 GGAACATTTGAAGTTGAAGATCAAAATTTGAAGCAGCCAGACAAATTTTCAAAAATGGGATTT 1879
QY 1741 GTGACACAACAAACCAATTTGCAATTTGGGCTGGTGCATATGAGGGGTAGCTAACCTCAATG 1800
Db 1880 GTGACACAACAAACCAATTTGCAATTTGGGCTGGTGCATATGAGGGGTAGCTAACCTCAATG 1939
QY 1801 GTCTCTGGGATCGGGAAGTGGCGTGTTCAAAGTGTGGAATAGCGCTGGCGCTGTATCCCGG 1860
Db 1940 GTCTCTGGGATCGGGAAGTGGCGTGTTCAAAGTGTGGAATAGCGCTGGCGCTGTATCCCGG 1999
QY 1861 TGGGAGTACTATGATCTAGTGTACAAGAACGTTACATGCGGTCTCCCAACTCCAGAAAGAC 1920
Db 2000 TGGGAGTACTATGATCTAGTGTACAAGAACGTTACATGCGGTCTCCCAACTCCAGAAAGAC 2059
QY 1921 AACCTTCGACCAATTCAGAAATTTCAACAGTGCATGAGCAGAGCTGAAATTTTAAACAAGTT 1980
Db 2060 AACCTTCGACCAATTCAGAAATTTCAACAGTGCATGAGCAGAGCTGAAATTTTAAACAAGTT 2119
QY 1981 GAGTACCTCTCTTATTTATGGAACAGCAGATGATAACGTTTCACTTTTCAAGCTCAGCTCAG 2040
Db 2120 GAGTACCTCTCTTATTTATGGAACAGCAGATGATAACGTTTCACTTTTCAAGCTCAGCTCAG 2179
QY 2041 ATCTCCAAAGCCCTGGTGCATGTTGGAGTGGATTTCCAGGCAATGTGGTATACGTATGAA 2100
Db 2180 ATCTCCAAAGCCCTGGTGCATGTTGGAGTGGATTTCCAGGCAATGTGGTATACGTATGAA 2239
QY 2101 GACCATGGATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2160
Db 2240 GACCATGGATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2299
QY 2161 ATAAAAACAATGTTTCTCTTTACCT 2184
Db 2300 ATAAAAACAATGTTTCTCTTTACCT 2323
```



Db	1485	GTGTCATTCAGTAAAGAGCGAGTATTATTCAGCTGAGATGTTCCGGTCTCGGTCTGCC	1544
Qy	1321	CTCTATACTCTACACAGCAGCGTGAATGATAAAGGGCTGAGAGTCTCTGGAGACAATTC	1380
Db	1545	CTCTATACTCTACACAGCAGCGTGAATGATAAAGGGCTGAGAGTCTCTGGAGACAATTC	1604
Qy	1381	GCTTTGGATAAATGCTGCAGAAATCCAGATGCCCTCCAAAACCTGGATCTCATTT	1440
Db	1605	GCTTTGGATAAATGCTGCAGAAATCCAGATGCCCTCCAAAACCTGGATCTCATTT	1664
Qy	1441	TTGAATGAAAACAAATTTTGGTATCAGATGATCTTGGCTCTCAATTTTGATAAATCCAG	1500
Db	1665	TTGAATGAAAACAAATTTTGGTATCAGATGATCTTGGCTCTCAATTTTGATAAATCCAG	1724
Qy	1501	AAATATCCTCTATCTATTAGATGTTATGCAAGGCCCATGTAGTCAAAAAGCAGACACTGTC	1560
Db	1725	AAATATCCTCTATCTATTAGATGTTATGCAAGGCCCATGTAGTCAAAAAGCAGACACTGTC	1784
Qy	1561	TTCAGACTGAACCTGGGCCACTTACTCTTGCAGACACAGAAAACATTAATAGTAGCTAGCTT	1620
Db	1785	TTCAGACTGAACCTGGGCCACTTACTCTTGCAGACACAGAAAACATTAATAGTAGCTAGCTT	1844
Qy	1621	GATGGCAGAGGAAGTGTATTACCAAGGAGATAAGATCATGCAATGCAATCAACAGAACTG	1680
Db	1845	GATGGCAGAGGAAGTGTATTACCAAGGAGATAAGATCATGCAATCAACAGAACTG	1904
Qy	1681	GGAAACATTTGAAATGGAAGATCAAAATTTGAAGCAAGCCAGACAATTTTCAAAAATGGGATTT	1740
Db	1905	GGAAACATTTGAAATGGAAGATCAAAATTTGAAGCAAGCCAGACAATTTTCAAAAATGGGATTT	1964
Qy	1741	GTGGACAACAAACGAAATTTGCAATTTTGGGGCTGGTCAATATCGAGGGTACGTAACTCAATG	1800
Db	1965	GTGGACAACAAACGAAATTTGCAATTTTGGGGCTGGTCAATATCGAGGGTACGTAACTCAATG	2024
Qy	1801	GTCTCTGGGATCGGGAAGTGGCGTGTTCAAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG	1860
Db	2025	GTCTCTGGGATCGGGAAGTGGCGTGTTCAAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG	2084
Qy	1861	TGGGAGTACTATGACTCAGTGTACACAGAACGTTACATGCGGTCTCCAACTCAGAAAGAC	1920
Db	2085	TGGGAGTACTATGACTCAGTGTACACAGAACGTTACATGCGGTCTCCAACTCAGAAAGAC	2144
Qy	1921	AACCTTGACCAATACAGAAATTCAAAGTCAATGAGCAGAGCTGAAAATTTTAAACAAGTT	1980
Db	2145	AACCTTGACCAATACAGAAATTCAAAGTCAATGAGCAGAGCTGAAAATTTTAAACAAGTT	2204
Qy	1981	GAGTACCTCTTATTCATGGAACAGCAGATGATAAGTTCATTTGACGAGTCAGCTCAG	2040
Db	2205	GAGTACCTCTTATTCATGGAACAGCAGATGATAAGTTCATTTGACGAGTCAGCTCAG	2264
Qy	2041	ATCTCCAAAGCCCTGGTTCGATGTTGGAGTGGATTTCCAGGCATGTGGTATCTGATGAA	2100
Db	2265	ATCTCCAAAGCCCTGGTTCGATGTTGGAGTGGATTTCCAGGCATGTGGTATCTGATGAA	2324
Qy	2101	GACCATGGAATAGCTAGCAGCACAGCACCAACATATATATACCAATGAGCCACTTC	2160
Db	2325	GACCATGGAATAGCTAGCAGCACAGCACCAACATATATATACCAATGAGCCACTTC	2384
Qy	2161	ATAAAACAAATGTTCTCTTTACCT	2184
Db	2385	ATAAAACAAATGTTCTCTTTACCT	2408

RESULT 14

AD019397

ID ADO19397 standard; cDNA: 3445 bp.

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

AC ADO19397;

XX

DT 12-AUG-2004 (first entry)

100

DE Human PRO polynucleotide #164.

	Human; PRO; Gene; ss; immune related disorder; systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis; systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis; autoimmune haemolytic anaemia; autoimmune thrombocytopenia; thyroiditis; diabetes mellitus; renal disease; demyelinating disease; central nervous system; peripheral nervous system; demyelinating polyneuropathy; Guillain-Barre syndrome; chronic inflammatory demyelinating polyneuropathy.
XX	Homosapiens.
OS	WO2004043361-A2.
XN	27-MAY-2004.
XX	06-NOV-2003; 2003WO-US035268.
PF	08-NOV-2002; 2002US-0425235P.
PR	(GETH ) GENENTECH INC.
XX	Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
PI	Wood WI, Wu TD;
FI	WPI; 2004-420067/39.
DR	P-PSDB; ADO19398.
DR	Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO6388 useful for treating an immune related disorder such as systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or spondyloarthropathy.
PT	Claim 1; SEQ ID NO 340; 1731pp; English.
XX	The invention relates to human PRO polypeptides and the polynucleotides encoding them. The polypeptides and polynucleotides are useful for treating and diagnosing immune related disorders in mammals. The immune related disorders include systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis, systemic sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal disease, demyelinating diseases of the central or peripheral nervous system, demyelinating polyneuropathy, Guillain-Barre syndrome and chronic inflammatory demyelinating polyneuropathy. This sequence represents a human PRO polynucleotide of the invention.
CC	Sequence 3445 BP; 1085 A; 689 C; 721 G; 950 T; 0 U; 0 Other;
XX	Query Match 99.9%; Score 2182.4; DB 12; Length 3445;
SQ	Best Local Similarity 99.9%; Pred No. 0;
	Matches 2183; Conservative 0; Mismatches 1; Indels 0; Gaps 0
QY	1 AGTCGCAAACTTACACTCTAACGTATTACTTTAAAAAATCTTATAGACTGAAGTTATAC 60
Db	225 AGTCGCAAACTTACACTCTAACGTATTACTTTAAAAAATCTTATAGACTGAAGTTATAC 284
QY	61 TCCTTAAGATGGATTTCAGATCATGAATATCTCTACAAAAGAATAATATCTTGTTA 120
Db	285 TCCTTAAGATGGATTTCAGATCATGAATATCTCTACAAAAGAATAATATCTTGTTA 344
QY	121 TTCAAATGCTCAATATGAAAACAGCTCAGTTTTCTTGGAGACACAGTACATTTGATCAGTTT 180
Db	345 TTCAAATGCTCAATATGAAAACAGCTCAGTTTTCTTGGAGACACAGTACATTTGATCAGTTT 404
QY	181 GGACATTTCTATCAATGATTTATTCATATCTCCTGATGGCAGTTTATTTCTCTTAGAATAC 240
Db	405 GGACATTTCTATCAATGATTTATTCATATCTCCTGATGGCAGTTTATTTCTCTTAGAATAC 464
QY	241 AACCTAGTGAAGCAATGGAGCATTCCTCACAGCTTCATATGACATTTATGATTTAAT 300

Db 465 AACTACGTGAAGCAATGGAGGCAATTCCTTACACAGCTTCATATGACATTTATGATTTAAAT 524  
Qy 301 AAAAGGACGCTGATTACAGAGAGAGGATTCAAACAAACAACACACAGTGGGTCAATGGTCA 360  
Db 525 AAAAGGACGCTGATTACAGAGAGAGGATTCAAACAAACAACACACAGTGGGTCAATGGTCA 584  
Qy 361 CCAAGTGGGTCAATAATTTGCAATATGTTTGGAAACAATGACATTTATGTTTAAATTTGAACCA 420  
Db 585 CCAAGTGGGTCAATAATTTGCAATATGTTTGGAAACAATGACATTTATGTTTAAATTTGAACCA 644  
Qy 421 AATTTACCAAGTTTACAGAAATCACATGGACGGGAAAGAGATATATATATATATGAAATA 480  
Db 645 AATTTACCAAGTTTACAGAAATCACATGGACGGGAAAGAGATATATATATATATGAAATA 704  
Qy 481 ACTGACTGGGTTTATGAAGAGAAAGTCTTTCAGTGCCTTACTCTGCTCTGTTGGTGGTCTCCA 540  
Db 705 ACTGACTGGGTTTATGAAGAGAAAGTCTTTCAGTGCCTTACTCTGCTCTGTTGGTGGTCTCCA 764  
Qy 541 AACGGCACTTTTATAGCATATGCCCCAATTTTAAAGACACAGAGTCCCACTTATTTGAATAC 600  
Db 765 AACGGCACTTTTATAGCATATGCCCCAATTTTAAAGACACAGAGTCCCACTTATTTGAATAC 824  
Qy 601 TCCTTCTACTCTGATGAGTCACTGACAGTACCCAAAGACTGTACGGGTTCCATATCCAAAG 660  
Db 825 TCCTTCTACTCTGATGAGTCACTGACAGTACCCAAAGACTGTACGGGTTCCATATCCAAAG 884  
Qy 661 GCAGGAGCTGTGAATCCAACTGTAAAGTCTTTGTTGTTGTTAAATAACAGACTCTCTCAGCTCA 720  
Db 885 GCAGGAGCTGTGAATCCAACTGTAAAGTCTTTGTTGTTGTTAAATAACAGACTCTCTCAGCTCA 944  
Qy 721 GTACCAATGCAACTTCCATACAAATCACTGCTCTGCTTCTATGTTGATAGGGATCAC 780  
Db 945 GTACCAATGCAACTTCCATACAAATCACTGCTCTGCTTCTATGTTGATAGGGATCAC 1004  
Qy 781 TACTTGTGTGATGTGATGAGTGGGCAACAAGAAAGAAATTTCTTTGAGTGGCTCAGGAGG 840  
Db 1005 TACTTGTGTGATGTGATGAGTGGGCAACAAGAAAGAAATTTCTTTGAGTGGCTCAGGAGG 1064  
Qy 841 ATTCAGAACTATTCGGTCAATGGAATTTGTGACTATGATGAATCCAGTGGAGAGATGGAAC 900  
Db 1065 ATTCAGAACTATTCGGTCAATGGAATTTGTGACTATGATGAATCCAGTGGAGAGATGGAAC 1124  
Qy 901 TGCCTTAGTGGCAGGCAACACATTTGAATGAGTACTCTGCTGGCTGGGTGGAGAGATTTAGG 960  
Db 1125 TGCCTTAGTGGCAGGCAACACATTTGAATGAGTACTCTGCTGGCTGGGTGGAGAGATTTAGG 1184  
Qy 961 CCTTCAGAACTCATTTTACCCTTGTATGTTGATGATGCTTCTACAGATCATCAGCAATGAA 1020  
Db 1185 CCTTCAGAACTCATTTTACCCTTGTATGTTGATGATGCTTCTACAGATCATCAGCAATGAA 1244  
Qy 1021 GAAGGTTACAGACACATTTGCTATTTTCCAAATAGATAAAGAACTGCAATTTATTTACA 1080  
Db 1245 GAAGGTTACAGACACATTTGCTATTTTCCAAATAGATAAAGAACTGCAATTTATTTACA 1304  
Qy 1081 AAAGGCACTGGGAAAGTCACTGGGATGAGAGTCTTAACAGATGATTTCTATCTACATTT 1140  
Db 1305 AAAGGCACTGGGAAAGTCACTGGGATGAGAGTCTTAACAGATGATTTCTATCTACATTT 1364  
Qy 1141 AGTAATGAATATATAAGGAATGCCAGGAGGAAGAAATCTTTTATAAAATCCAACTTATTTGAC 1200  
Db 1365 AGTAATGAATATATAAGGAATGCCAGGAGGAAGAAATCTTTTATAAAATCCAACTTATTTGAC 1424  
Qy 1201 TATACAAAAGTGACATGCTCAGTGTGAGCTGAATCCGGAAAGGTTGTCAGTACTATTTCT 1260  
Db 1425 TATACAAAAGTGACATGCTCAGTGTGAGCTGAATCCGGAAAGGTTGTCAGTACTATTTCT 1484  
Qy 1261 GTGCTCAATCAGTAAAGGCGAGTATATCAGCTCAGATGTTCCGGTCTGCTGTCGCC 1320  
Db 1485 GTGCTCAATCAGTAAAGGCGAGTATATCAGCTCAGATGTTCCGGTCTGCTGTCGCC 1544  
Qy 1321 CTCTATACCTACACAGCAGCGTGAATGATTAAGGGCTGAGAGTCTCTGGAAACAATTTCA 1380  
Db 1545 CTCTATACCTACACAGCAGCGTGAATGATTAAGGGCTGAGAGTCTCTGGAAACAATTTCA 1604

Qy 1381 GCTTTGGATATAAATGCTCCAGAAATGTCAGATGCCCTCCAAAAAATCGACTTTCATTATT 1440  
Db 1605 GCTTTGGATATAAATGCTCCAGAAATGTCAGATGCCCTCCAAAAAATCGACTTTCATTATT 1664  
Qy 1441 TTGAATGAAACAAAAATTTTGGTATCAGATGATCTTTGGCTCTCTCAATTTTGAATAAATCCAAG 1500  
Db 1665 TTGAATGAAACAAAAATTTTGGTATCAGATGATCTTTGGCTCTCTCAATTTTGAATAAATCCAAG 1724  
Qy 1501 AAATATCTCTACTATATAGATGTGAGGCCCCATGTAGTCAAAAAGCAGACACTGTC 1560  
Db 1725 AAATATCTCTACTATATAGATGTGAGGCCCCATGTAGTCAAAAAGCAGACACTGTC 1784  
Qy 1561 TTCAGACTGAATCGGCCACTTACCTTGCAGAGCAGAAAAACATTTATAGTACTAGCTTT 1620  
Db 1785 TTCAGACTGAATCGGCCACTTACCTTGCAGAGCAGAAAAACATTTATAGTACTAGCTTT 1844  
Qy 1621 GATGCGCAGAGGAAGTGGTTTACCAAGGAGATGAAGATCATGCAATCAACAGAGAGACTG 1680  
Db 1845 GATGCGCAGAGGAAGTGGTTTACCAAGGAGATGAAGATCATGCAATCAACAGAGAGACTG 1904  
Qy 1681 GGAAATTTGAAGTTGAAGATCAAATTTGAAGCAGCAGACAATTTTCAAAAATGGGATTT 1740  
Db 1905 GGAAATTTGAAGTTGAAGATCAAATTTGAAGCAGCAGACAATTTTCAAAAATGGGATTT 1964  
Qy 1741 GTGGACAAACAAGAAATTTGGGCTGCTCATATGAGGGGTAGCTAACTCAATG 1800  
Db 1965 GTGGACAAACAAGAAATTTGGGCTGCTCATATGAGGGGTAGCTAACTCAATG 2024  
Qy 1801 GTCTTGGGATCGGGAAGTGGGCTGTTCAAGTGTGGAATAGCCGTGGCTGTATCCCCG 1860  
Db 2025 GTCTTGGGATCGGGAAGTGGGCTGTTCAAGTGTGGAATAGCCGTGGCTGTATCCCCG 2084  
Qy 1861 TGGGAGTACTATGACTCAGTGTACAGAAAGTACAGTGGTCTCCAACTCCAGAGAC 1920  
Db 2085 TGGGAGTACTATGACTCAGTGTACAGAAAGTACAGTGGTCTCCAACTCCAGAGAC 2144  
Qy 1921 AACCTTGACCAATACAGAAATTTCAACAGTCAAGCAGAGCTGAAAAATTTTAAACAAGTT 1980  
Db 2145 AACCTTGACCAATACAGAAATTTCAACAGTCAAGCAGAGCTGAAAAATTTTAAACAAGTT 2204  
Qy 1981 GAGTACCTCTTATTTATGGAACAGCAGATGATAAAGTTTCACTTTCAGCAGTCAGCTCAG 2040  
Db 2205 GAGTACCTCTTATTTATGGAACAGCAGATGATAAAGTTTCACTTTCCAGCAGTCAGCTCAG 2264  
Qy 2041 ATCTCCAAAGCCCTGGTGTGATGTTGGAGTGGATTTTCCAGGCAATGTGGTATGATGAA 2100  
Db 2265 ATCTCCAAAGCCCTGGTGTGATGTTGGAGTGGATTTTCCAGGCAATGTGGTATGATGAA 2324  
Qy 2101 GACCATGGAATAGCTAGCAGCAGACACCAACATATATATACCAATGAGCCACTTTC 2160  
Db 2325 GACCATGGAATAGCTAGCAGCAGACACCAACATATATATACCAATGAGCCACTTTC 2384  
Qy 2161 ATAAAAACAATGTTTCTCTTTACCT 2184  
Db 2385 ATAAAAACAATGTTTCTCTTTACCT 2408

RESULT 15

ADO71613

ID ADO71613 standard; DNA; 3445 BP.

XX ADO71613;

XX ADO71613;

DT 26-AUG-2004 (first entry)

XX Nucleotide sequence of a human CD26 protein.

XX CD26; chemotherapeutic; radiotherapeutic; cancer; cell growth;

XX dipeptidyl peptidase IV; DPPIV; topoisomerase II inhibitor;

XX cell cycle arrest; tumour; tumour necrosis; immune response; human; gene;

XX SB.

XX



OS Homo sapiens:  
FH Key Location/Qualifiers  
FT CDS 111..2411  
FT /\*tag= a  
FT /product= "CD26"  
FN WO2004045497-A2.  
XX 03-JUN-2004.  
XX  
XX  
XX 15-MAY-2003; 2003WO-US015499.  
XX  
XX 17-MAY-2002; 2002US-0381606P.  
XX (TEXA ) UNIV TEXAS SYSTEM.  
XX  
XX Dang NH, Morimoto C;  
XX  
XX WPI; 2004-420511/39.  
XX P-PSDB; ADO71612.  
XX  
XX Use of a CD26 composition, and a chemotherapeutic and/or a  
XX radiotherapeutic agent for e.g. inhibiting the cell growth, inducing cell  
XX cycle arrest, killing a cancer cell, treating cancer, or inducing tumor  
XX regression or tumor necrosis.  
XX  
XX Claim 14; Page 153-157; 182pp; English.  
XX  
XX The specification describes a CD26 composition which, in conjunction with  
XX chemotherapeutic or radiotherapeutic agents, is used for the treatment  
XX and prevention of cancers. Expression of CD26 enhances the sensitivity of  
XX the cancer cell to the chemotherapeutic or radiotherapeutic agent. CD26  
XX is a dipeptidyl peptidase IV (DPP-IV). The chemotherapeutic agent is a  
XX topoisomerase II inhibitor. The CD26 composition of the invention is  
XX useful for inhibiting the growth of a cell, inducing cell cycle arrest in  
XX a cell, killing a cancer cell, potentiating the effect of a  
XX chemotherapeutic agent and/or a radiotherapeutic agent on a tumor cell,  
XX inducing or enhancing apoptosis of a cancer cell, treating cancer, or  
XX furthering tumor regression or tumor necrosis. The CD26 composition is  
XX activating an antigen-presenting cell, or for potentiating immune  
XX responses of an animal. The present sequence encodes a CD26 protein, and  
XX is used in vectors which are used to produce compositions of the  
XX invention.  
XX  
XX Sequence 3445 BP; 1085 A; 689 C; 721 G; 950 T; 0 U; 0 Other;  
SQ

Query Match 99.9%; Score 2182.4; DB 12; Length 3445;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2183; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AGTCGCAAACTTACACTCTAAGTATCTTAAATACTTAAATACTTATAGCTGAGTTATAC 60  
Db 225 AGTCGCAAACTTACACTCTAAGTATCTTAAATACTTAAATACTTATAGCTGAGTTATAC 284  
Qy 61 TCCTTAAGATGGATTTCAGATCATGAATATCTTACAAACAAGAAATAATATCTTGTA 120  
Db 285 TCCTTAAGATGGATTTCAGATCATGAATATCTTACAAACAAGAAATAATATCTTGTA 344  
Qy 121 TTCAATGCTGAATATGGAACAGCTCAGTTTTCTTGGAGAACAGTACATATTTGATGAGTTT 180  
Db 345 TTCAATGCTGAATATGGAACAGCTCAGTTTTCTTGGAGAACAGTACATATTTGATGAGTTT 404  
Qy 181 GGACATTTCTATCAATGATTAATTCATATCTCTGATGGGAGTTTATCTTCTTAGAATAC 240  
Db 405 GGACATTTCTATCAATGATTAATTCATATCTCTGATGGGAGTTTATCTTCTTAGAATAC 464  
Qy 241 AACTACGTGAAGCAATGAGGCAATTCCTTACACAGCTTCATATGACATTTATGATTTAAAT 300  
Db 465 AACTACGTGAAGCAATGAGGCAATTCCTTACACAGCTTCATATGACATTTATGATTTAAAT 524  
Qy 301 AAAAGGCAGCTGATTACAGAGAGAGGATTCCAAACAACAACAACAGTGGGTCACTAGTCA 360

Db 525 AAAAGGCAGCTGATTACAGAGAGAGGATTCCAAACAACAACAACAGTGGGTCACTAGTCA 584  
Qy 361 CCAGTGGGTCAATAAATTGGGCATATGTTTGGACACATGACATTTATTTAAATTTAAACCA 420  
Db 585 CCAGTGGGTCAATAAATTGGGCATATGTTTGGACACATGACATTTATTTAAATTTAAACCA 644  
Qy 421 AATTTACCAAGTTACAGAAATCACATGACGCGGGAAGAAAGATATAATATAATGAATA 480  
Db 645 AATTTACCAAGTTACAGAAATCACATGACGCGGGAAGAAAGATATAATATAATGAATA 704  
Qy 481 ACTGACTGGGTTTATGAAGAGAGTCTTCAAGTGCCTACTCTGCTCTGTGTGGTCTCCA 540  
Db 705 ACTGACTGGGTTTATGAAGAGAGTCTTCAAGTGCCTACTCTGCTCTGTGTGGTCTCCA 764  
Qy 541 AACGGCACTTTTATGACATATGCCCAATTTAAACGACACAGAGTCCCACTTATTTGAATAC 600  
Db 765 AACGGCACTTTTATGACATATGCCCAATTTAAACGACACAGAGTCCCACTTATTTGAATAC 824  
Qy 601 TCCTTTACTCTGATGAGTCACTGACGATCCCAAGACTGTACGGGTTCCATATCCAAAG 660  
Db 825 TCCTTTACTCTGATGAGTCACTGACGATCCCAAGACTGTACGGGTTCCATATCCAAAG 884  
Qy 661 GCAGGAGCTGGAATCCAACTGTAAAGTTCTTTGTGTAAATACAGACTCTCTCAGCTCA 720  
Db 885 GCAGGAGCTGGAATCCAACTGTAAAGTTCTTTGTGTAAATACAGACTCTCTCAGCTCA 944  
Qy 721 GTCCCAATGCACTTCCATACAAATCACTGCTCTGCTCTCTATGTTGATAGGGATCAC 780  
Db 945 GTCCCAATGCACTTCCATACAAATCACTGCTCTGCTCTCTATGTTGATAGGGATCAC 1004  
Qy 781 TACTTGTGTGATGACATGGGCAACAAGAAAGATTTCTTTGACGTGCTCAGGAGG 840  
Db 1005 TACTTGTGTGATGACATGGGCAACAAGAAAGATTTCTTTGACGTGCTCAGGAGG 1064  
Qy 841 ATTCCAGAACTTATTCGGTCAATGATATTTGTGACTATGATGAATCCAGTGGAGATGGAAC 900  
Db 1065 ATTCCAGAACTTATTCGGTCAATGATATTTGTGACTATGATGAATCCAGTGGAGATGGAAC 1124  
Qy 901 TGCTTACTGCGACGGCAACATTTGAATGAGTACTACTGCTGGTGGTGGAGATTTAGG 960  
Db 1125 TGCTTACTGCGACGGCAACATTTGAATGAGTACTACTGCTGGTGGTGGAGATTTAGG 1184  
Qy 961 CCTTCAGAACTTATTTTACCTTTGATGTAATAGCTTTCTACAAGATCATCAGCAATGAA 1020  
Db 1185 CCTTCAGAACTTATTTTACCTTTGATGTAATAGCTTTCTACAAGATCATCAGCAATGAA 1244  
Qy 1021 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATATAAAGACTGACATTTATTACA 1080  
Db 1245 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATATAAAGACTGACATTTATTACA 1304  
Qy 1081 AAAGGCACCTGGGAAGTCACTGGGATAGAGCTCTTAACCGATGATTATCTATCTACATT 1140  
Db 1305 AAAGGCACCTGGGAAGTCACTGGGATAGAGCTCTTAACCGATGATTATCTATCTACATT 1364  
Qy 1141 AGTAATGAATATAAAGGAATGCCAGGAGGAGGAATCTTTATAAAATCCAACTTATTGAC 1200  
Db 1365 AGTAATGAATATAAAGGAATGCCAGGAGGAGGAATCTTTATAAAATCCAACTTATTGAC 1424  
Qy 1201 TATACAAAAGTACATGCCCTCAGTTGTGAGCTGAATCCGGAAGAGGTGTCAGTACTATTCT 1260  
Db 1425 TATACAAAAGTACATGCCCTCAGTTGTGAGCTGAATCCGGAAGAGGTGTCAGTACTATTCT 1484  
Qy 1261 GTGTCAATCAGTAAGAGGCGAAGATTTATCAGCTGAGATGTTCCGGTCTCGGTCTGCC 1320  
Db 1485 GTGTCAATCAGTAAGAGGCGAAGATTTATCAGCTGAGATGTTCCGGTCTCGGTCTGCC 1544  
Qy 1321 CTCTATCTCTACACAGCAGCTGAATGATAAAGGGCTGAGAGTCTCGGGAAGACAAATTCA 1380  
Db 1545 CTCTATCTCTACACAGCAGCTGAATGATAAAGGGCTGAGAGTCTCGGGAAGACAAATTCA 1604  
Qy 1381 GCTTTGGATAAAATGCTGCAAGATGTCCAGATGCCCTCCAAAAATCTGGACTTCAATTAT 1440





GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 17, 2006, 04:24:53 ; Search time 11140 Seconds  
(without alignments)  
11144.174 Million cell updates/sec

Title: US-10-659-055-2  
Perfect score: 2184  
Sequence: 1 agtcgcaaaccttacactct.....aacaaatgttcttcttacct 2184

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues  
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_in.\*
- 3: gb\_env.\*
- 4: gb\_on.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pr.\*
- 9: gb\_ro.\*
- 10: gb\_sts.\*
- 11: gb\_sy.\*
- 12: gb\_un.\*
- 13: gb\_vi.\*
- 14: gb\_htg.\*
- 15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2184	100.0	2211	6	CQ815921 Sequence
2	2184	100.0	3407	6	CQ963950 Sequence
3	2184	100.0	3407	6	AR349829 Sequence
4	2184	100.0	3407	6	AR653630 Sequence
5	2184	100.0	3407	8	HSFCHDP7
6	2182.4	99.9	2301	6	CQ772397 Sequence
7	2182.4	99.9	2431	8	BC013329 Homo sapi
8	2182.4	99.9	3411	6	CQ727953 Sequence
9	2182.4	99.9	3445	6	CQ776356 Sequence
10	2182.4	99.9	3445	6	CQ981570 Sequence
11	2182.4	99.9	3445	6	CS031460 Sequence
12	2182.4	99.9	3445	6	CS032879 Sequence
13	2182.4	99.9	3445	6	CS040412 Sequence
14	2182.4	99.9	3445	6	CS041831 Sequence
15	2182.4	99.9	3913	8	BC065265 Homo sapi
16	2182.4	99.9	3948	6	CS134667 Sequence
17	2180.8	99.9	2924	6	AR380481 Sequence
18	2180.8	99.9	2924	8	HUMCD26A M74777 Human dipep

19	2177.6	99.7	2375	8	HUMDIPBPV
20	1852.8	84.8	2573	4	AB023952
21	1844.8	84.5	2727	4	AF461806
22	1843.2	84.4	2298	4	AY056834
23	1827.8	83.7	2400	4	AY198323
24	1677.6	76.8	3316	6	CQ777283
25	1677.6	76.8	3316	9	MMDPP4
26	1677.6	76.8	3482	9	BC022183
27	1657.2	75.9	3214	9	RATGPMBCD
28	1656.2	75.8	4835	6	AX401894
29	1656.2	75.8	4835	9	RATDPP
30	1009.2	46.2	3203	5	AJ720903
31	923.2	42.3	4250	5	AB158225
32	923.2	42.3	4254	5	AB158224
33	762.4	34.9	2834	5	CR926279
34	761.6	34.9	3376	5	BC082401
35	754.2	34.5	2366	8	HSU76833
36	752.6	34.5	2648	8	BC026250
37	751	34.4	2814	6	AX329990
38	751	34.4	2814	6	AX330925
39	751	34.4	2814	6	AX333666
40	751	34.4	2814	6	AX335485
41	751	34.4	2814	6	AX578046
42	751	34.4	2814	8	HSU09278
43	751	34.4	2815	6	AR078938
44	751	34.4	2815	6	I32308
45	751	34.4	2815	6	AR633310

ALIGNMENTS

RESULT 1  
CQ815921  
LOCUS CQ815921 2211 bp DNA linear PAT 03-JUN-2004  
DEFINITION Sequence 1 from Patent EPI422293.  
ACCESSION CQ815921  
VERSION CQ815921.1 GI:48144442  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens

REFERENCE 1  
AUTHORS Hennig, M., Loeffler, B. M. and Thoma, R.  
TITLE Crystal structure of dipeptidyl peptidase IV  
JOURNAL Patent: EP 1422293-A 1 26-MAY-2004;  
F. HOFFMANN-LA ROCHE AG (CH)  
FEATURES  
source  
1. .2211  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 100.0%; Score 2184; DB 6; Length 2211;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	AGTCGCAAACTTACACTCTAACTGATTACTTAAAAAATCTTATAGACGAACTTATAC	60
Db	25	AGTCGCAAACTTACACTCTAACTGATTACTTAAAAAATCTTATAGACGAACTTATAC	84
Qy	61	TCCTTAAGATGGATTTTCAGATCATGATATCTCTACAAACAGAAATATATCTTGGTA	120
Db	85	TCCTTAAGATGGATTTTCAGATCATGATATCTCTACAAACAGAAATATATCTTGGTA	144
Qy	121	TTCAATCTCAATATGGAACAGCTAGTTTCTTTGGAGAACAGTACATCTTGGAGTTT	180
Db	145	TTCAATCTCAATATGGAACAGCTAGTTTCTTTGGAGAACAGTACATCTTGGAGTTT	204
Qy	181	GGACATCTTATCAATGATTATCAATATCTCTCTGATGGCGAGTTTATTTCTTTAGAAATAC	240



Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

REFERENCE

1  
AUTHORS  
TITLE  
Diagnosics and therapeutics for diseases associated with  
dipeptidylpeptidase iv (dpp4)  
JOURNAL  
Patent: WO 2004104216-A 1 02-DEC-2004;  
Bayer HealthCare AG (DE)  
FEATURES  
Location/Qualifiers  
source  
1. .3407

/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 2184; DB 6; Length 3407;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 AGTCGCAAACTTACACTCTAACTGATTAATCTTAAATAAATACTTATAGACTGAAGTTATAC 60  
Db |  
Qy 190 AGTCGCAAACTTACACTCTAACTGATTAATCTTAAATAAATACTTATAGACTGAAGTTATAC 249  
Db |  
Qy 61 TCCTTAAGATGGATTCAGATCATGATATCTCTACAAACAGAAATATATCTTGGTA 120  
Db |  
Qy 250 TCCTTAAGATGGATTCAGATCATGATATCTCTACAAACAGAAATATATCTTGGTA 309  
Db |  
Qy 121 TTCAATGCTGAATATGAAACAGCTCAGTTTCTTGGAGACAGTACATTTGATGAGTTT 180  
Db |  
Qy 310 TTCAATGCTGAATATGAAACAGCTCAGTTTCTTGGAGACAGTACATTTGATGAGTTT 369  
Db |  
Qy 181 GGACATCTATCAATGATTAATCAATATCTCTGATGGCAGTTTATCTCTTGAATATAC 240  
Db |  
Qy 370 GGACATCTATCAATGATTAATCAATATCTCTGATGGCAGTTTATCTCTTGAATATAC 429  
Db |  
Qy 241 AACTACGTGAAGCAATGGAGGCATTCCTACAGCTTCATATGATGATTAAT 300  
Db |  
Qy 430 AACTACGTGAAGCAATGGAGGCATTCCTACAGCTTCATATGATGATTAAT 489  
Db |  
Qy 301 AAAAGGCGAGCTGATTAACAGAGAGAGGATTCCTCAACACACACAGTGGGTCAATGTCA 360  
Db |  
Qy 490 AAAAGGCGAGCTGATTAACAGAGAGAGGATTCCTCAACACACACAGTGGGTCAATGTCA 549  
Db |  
Qy 361 CCAGTGGGTCTATAAATGGCATATGTTTGGACCAATGACATTTATGTTAAATGAAACCA 420  
Db |  
Qy 550 CCAGTGGGTCTATAAATGGCATATGTTTGGACCAATGACATTTATGTTAAATGAAACCA 609  
Db |  
Qy 421 AATTACCAAGTTACAGATCAGATCAGTGGGGAAGAGATATATATATATATGGAATA 480  
Db |  
Qy 610 AATTACCAAGTTACAGATCAGATCAGTGGGGAAGAGATATATATATATATGGAATA 669  
Db |  
Qy 481 ACTGACTGGGTTTATGAGGAGAGTCTTCAAGTGCCTACTCTGCTGTGGTGGTCTCCA 540  
Db |  
Qy 670 ACTGACTGGGTTTATGAGGAGAGTCTTCAAGTGCCTACTCTGCTGTGGTGGTCTCCA 729  
Db |  
Qy 541 AACGGCACTTTTTTATGATATGCCCATAATTAACGACACAGAGTCCCACTTATGGAATAC 600  
Db |  
Qy 730 AACGGCACTTTTTTATGATATGCCCATAATTAACGACACAGAGTCCCACTTATGGAATAC 789  
Db |  
Qy 601 TCCTTCTACTCTGATGATGATCTGATGATCCCAAGACTGTAACGGGTTCCATATCCAAAG 660  
Db |  
Qy 790 TCCTTCTACTCTGATGATGATCTGATGATCCCAAGACTGTAACGGGTTCCATATCCAAAG 849  
Db |  
Qy 661 GCAGGAGCTGTAATCCAACTGTAAGTCTTCTGTTGTAATACAGACTCTCTCAGCTCA 720  
Db |  
Qy 850 GCAGGAGCTGTAATCCAACTGTAAGTCTTCTGTTGTAATACAGACTCTCTCAGCTCA 909  
Db |  
Qy 721 GTCAACCAATGCACTTCCATAAATCACTGCTCTCTCTCTATGTTGATAGGGGATCAC 780  
Db |  
Qy 910 GTCAACCAATGCACTTCCATAAATCACTGCTCTCTCTCTATGTTGATAGGGGATCAC 969  
Db |  
Qy 781 TACTTGTGTGATGTGACATGGGCAACACAGAAAGAAATTTCTTTGAGTGGCTCAGGAGG 840  
Db |

Db 970 TACTTGTGTGATGTGACATGGGCAACACAGAAAGAAATTTCTTTGCGATGGCTCAGGAGG 1029  
Qy |  
Db 841 ATTACAGAACTATTTCGGTCATGGATATTTGTGACTATGATGAATCCAGTGGAGATGGAAC 900  
Qy |  
Db 1030 ATTACAGAACTATTTCGGTCATGGATATTTGTGACTATGATGAATCCAGTGGAGATGGAAC 1089  
Qy |  
Db 901 TGCTTAGTGCAACGGCAACACATTTGAAATGAGTACTACTGGCTGGGTTGGAAGATTTAGG 960  
Qy |  
Db 1090 TGCTTAGTGCAACGGCAACACATTTGAAATGAGTACTACTGGCTGGGTTGGAAGATTTAGG 1149  
Qy |  
Db 961 CCTTCAGAACTCATTTTTCCTTGATGTTAGTCTTCTACAGATCATCAGCAATGAA 1020  
Qy |  
Db 1150 CCTTCAGAACTCATTTTTCCTTGATGTTAGTCTTCTACAGATCATCAGCAATGAA 1209  
Qy |  
Db 1021 GAAGGTTACAGACACATTTTCTAATTTCCAAATAGATAAAAGAACTGCACATTTATTACA 1080  
Qy |  
Db 1210 GAAGGTTACAGACACATTTTCTAATTTCCAAATAGATAAAAGAACTGCACATTTATTACA 1269  
Qy |  
Db 1081 AAAGGCACTCTGGGAAGTTCATCGGATAGAGCTCTTAACAGTGAATATCTATATACATTT 1140  
Qy |  
Db 1270 AAAGGCACTCTGGGAAGTTCATCGGATAGAGCTCTTAACAGTGAATATCTATATACATTT 1329  
Qy |  
Db 1141 AGTAATGAATATAAGGAATGCCAGGAGGAGGATCTTTTATAAATCCAACTTATTGAC 1200  
Qy |  
Db 1330 AGTAATGAATATAAGGAATGCCAGGAGGAGGATCTTTTATAAATCCAACTTATTGAC 1389  
Qy |  
Db 1201 TATACAAAAGTGCATGCTCAGTTGTGAGCTGAATCCGGAAGGTTGTCACTATTTCT 1260  
Qy |  
Db 1390 TATACAAAAGTGCATGCTCAGTTGTGAGCTGAATCCGGAAGGTTGTCACTATTTCT 1449  
Qy |  
Db 1261 GTGTCAATTCAGTAAAGGCGAAGTATTATCAGCTGAGATGTTCCGGTCTCTGGTCTGCC 1320  
Qy |  
Db 1450 GTGTCAATTCAGTAAAGGCGAAGTATTATCAGCTGAGATGTTCCGGTCTCTGGTCTGCC 1509  
Qy |  
Db 1321 CTCTATCTCTACACAGCAGCGTGAATGAAGGCTGAGCTCTGGAAGACAAATTTCA 1380  
Qy |  
Db 1510 CTCTATCTCTACACAGCAGCGTGAATGAAGGCTGAGCTCTGGAAGACAAATTTCA 1569  
Qy |  
Db 1381 GCTTTGATGATAAATGCTGCAGAAATGTCAGATGCTCCAGTGCCTTCCAAAATCTCATTTAT 1440  
Qy |  
Db 1570 GCTTTGATGATAAATGCTGCAGAAATGTCAGATGCTCCAGTGCCTTCCAAAATCTCATTTAT 1629  
Qy |  
Db 1441 TTGAATGAAAACAAAATTTTGGTATCAGATGATCTTGCTCTCTCATTTTGTGATAAATCCAAAG 1500  
Qy |  
Db 1630 TTGAATGAAAACAAAATTTTGGTATCAGATGATCTTGCTCTCTCATTTTGTGATAAATCCAAAG 1689  
Qy |  
Db 1501 AAATATCTCTCTATTTAGATGTTAGTGCAGGCCCCATGTTAGTCAAAAGCAGACACTGTC 1560  
Qy |  
Db 1690 AAATATCTCTCTATTTAGATGTTAGTGCAGGCCCCATGTTAGTCAAAAGCAGACACTGTC 1749  
Qy |  
Db 1561 TTCAAGTCAAACTGGGCACTTACCTTGCAGACACAGAAACATTTATAGTAGCTAGCTTTT 1620  
Qy |  
Db 1750 TTCAAGTCAAACTGGGCACTTACCTTGCAGACACAGAAACATTTATAGTAGCTAGCTTTT 1809  
Qy |  
Db 1621 GATGGCAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGCAATGCAATCAACAGAGACTG 1680  
Qy |  
Db 1810 GATGGCAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGCAATGCAATCAACAGAGACTG 1869  
Qy |  
Db 1681 GGAACATTTGCAAGTTGAGATCAATTTGAAGCAGCCAGACAAATTTTCAAAAATGGGATTT 1740  
Qy |  
Db 1870 GGAACATTTGCAAGTTGAGATCAATTTGAAGCAGCCAGACAAATTTTCAAAAATGGGATTT 1929  
Qy |  
Db 1741 GTGCAACAAAACAAATTTGGGCTGGTTCATATGAGGAGTACGTAACCTCAATG 1800  
Qy |  
Db 1930 GTGCAACAAAACAAATTTGGGCTGGTTCATATGAGGAGTACGTAACCTCAATG 1889  
Qy |  
Db 1801 GTCTCTGGGATCGGGAAGTGGCTGTTCAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG 1860  
Qy |  
Db 1990 GTCTCTGGGATCGGGAAGTGGCTGTTCAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG 2049  
Qy |  
Db 1861 TGGGAGTACTATGACTCAGTGTACACAGACGTTTACATGGGCTCTCCCACTCCAGAGAC 1920  
Qy |  
Db 2050 TGGGAGTACTATGACTCAGTGTACACAGAACGTTTACATGGGCTCTCCCACTCCAGAGAC 2109  
Qy |

QY 1921 AACCTTGGACATTACAGAAATTCACAGCTCATGAGCAGAGCTGAAAAATTTTAAACAAGTT 1980  
DB 2110 AACCTTGACCATACAGAAATTCACAGCTCATGAGCAGAGCTGAAAAATTTTAAACAAGTT 2169  
QY 1981 GAGTACCTCTTATTCATGGAACAGCAGATGATAACGTTCACTTTTCAGCAGTCAGCTCAG 2040  
DB 2170 GAGTACCTCTTATTCATGGAACAGCAGATGATAACGTTCACTTTTCAGCAGTCAGCTCAG 2229  
QY 2041 ATCTCAAAAGCCCTGGTCGATGTTGGAGTGGATTTCCAGGCAATGTGGTATATCTGATGAA 2100  
DB 2230 ATCTCAAAAGCCCTGGTCGATGTTGGAGTGGATTTCCAGGCAATGTGGTATATCTGATGAA 2289  
QY 2101 GACCATGGAATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2160  
DB 2290 GACCATGGAATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2349  
QY 2161 ATAAAAACAATGTTCTCTTTACCT 2184  
DB 2350 ATAAAAACAATGTTCTCTTTACCT 2373  
  
RESULT 3  
LOCUS AR349829 3407 bp DNA linear PAT 17-AUG-2003  
DEFINITION Sequence 5 from patent US 6586198.  
ACCESSION AR349829  
VERSION AR349829.1 GI:33750723  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
  
REFERENCE  
1 (bases 1 to 3407)  
AUTHORS Brown, N.J.  
TITLE Method of identifying susceptibility to angiotensin converting  
enzyme inhibito- and vasopeptidase-inhibitor-associated angioedema  
JOURNAL Patent: US 6586198-A 5 01-JUL-2003;  
Vanderbilt University; Nashville, TN  
  
FEATURES  
source  
1..3407  
/organism="unknown"  
/mol\_type="genomic DNA"  
  
ORIGIN  
  
Query Match 100.0%; Score 2184; DB 6; Length 3407;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2184; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 AGTCGCAAACTTACACTCTAACTGATTACTTTAAAAATACTTATAGACTGAAGTTATAC 60  
DB 190 AGTCGCAAACTTACACTCTAACTGATTACTTTAAAAATACTTATAGACTGAAGTTATAC 249  
QY 61 TCCTTAAGATGGATTTCAGATCATGAATATCTCTACAAACAAGAAATAATATCTGGTA 120  
DB 250 TCCTTAAGATGGATTTCAGATCATGAATATCTCTACAAACAAGAAATAATATCTGGTA 309  
QY 121 TTCATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTGATGAGTTT 180  
DB 310 TTCATGCTGAATATGGAACAGCTCAGTTTCTTGGAGAACAGTACATTGATGAGTTT 369  
QY 181 GGACATCTCATCAATGATTATTCAATATCTCTGATGGSCAGTTTATCTCTTAGAATAC 240  
DB 370 GGACATCTCATCAATGATTATTCAATATCTCTGATGGSCAGTTTATCTCTTAGAATAC 429  
QY 241 AACTACGTGAAGCAATGGAGGCAATCTTACACAGCTTCATATGACATTTTATGATTTAAAT 300  
DB 430 AACTACGTGAAGCAATGGAGGCAATCTTACACAGCTTCATATGACATTTTATGATTTAAAT 489  
QY 301 AAAAGGCAGCTGATTACAGAGAGAGGATTCCAAACACACACAGTGGGTGACATGGTCA 360  
DB 490 AAAAGGCAGCTGATTACAGAGAGAGGATTCCAAACACACACAGTGGGTGACATGGTCA 549  
QY 361 CCAGTGGGTGATATAATTTGGCATATGTTTGGAAACAATGACATTTATGTTAAAAATGGAACCA 420

DB 550 CCAGTGGGTGATATAATTTGGCATATGTTTGGAAACAATGACATTTATGTTAAAAATGGAACCA 609  
QY 421 AATTATCCAAAGTTACAGATCAGATGAGCGGGGAAGAGATATATATATATATGGAATA 480  
DB 610 AATTATCCAAAGTTACAGATCAGATGAGCGGGGAAGAGATATATATATATGGAATA 669  
QY 481 ACTGACTGGGTTTATGAAGAGAGAGTCTTCAGTGCCTACTCTGCTCTCTGCTGCTGCTCCTCA 540  
DB 670 ACTGACTGGGTTTATGAAGAGAGAGTCTTCAGTGCCTACTCTGCTCTCTGCTGCTGCTCCTCA 729  
QY 541 AACGCGACTTTTATGACATATGCCCAATTTAACGACACAGAAAGTCCCACTTATTTGAATAC 600  
DB 730 AACGCGACTTTTATGACATATGCCCAATTTAACGACACAGAAAGTCCCACTTATTTGAATAC 789  
QY 601 TCCTTCTACTCTGATGAGTCACTGCACTGACCCAAAGACTGTACGGGTTCCATATCCAAAG 560  
DB 790 TCCTTCTACTCTGATGAGTCACTGCACTGACCCAAAGACTGTACGGGTTCCATATCCAAAG 849  
QY 661 GCAGGAGCTGTGAATCCCACTGTAAAGTCTTTGTTGTAATATACAGACTCTCTCAGCTCA 720  
DB 850 GCAGGAGCTGTGAATCCCACTGTAAAGTCTTTGTTGTAATATACAGACTCTCTCAGCTCA 909  
QY 721 CTCACCAATGCAACTTCCATACAAATCACTGCTCTGCTTCTATGTTGATAGGGAATCAC 780  
DB 910 CTCACCAATGCAACTTCCATACAAATCACTGCTCTGCTTCTATGTTGATAGGGAATCAC 969  
QY 781 TACTTGTGTGATGACATGCGGCAACACAGAAAGAAATTTCTTTCGAGTGGCTCAGGAGG 840  
DB 970 TACTTGTGTGATGACATGCGGCAACACAGAAAGAAATTTCTTTCGAGTGGCTCAGGAGG 1029  
QY 841 ATTCAGAACTATTCGGTCACTGATGATTTTGTGACTATGATGAATCCAGTGGAGATGGAAC 900  
DB 1030 ATTCAGAACTATTCGGTCACTGATGATTTTGTGACTATGATGAATCCAGTGGAGATGGAAC 1089  
QY 901 TGCTTAGTGGCACCGCAACACATTCGAAATGAGTACTACTGCTGGTGGTGGAGATTTAGG 960  
DB 1090 TGCTTAGTGGCACCGCAACACATTCGAAATGAGTACTACTGCTGGTGGTGGAGATTTAGG 1149  
QY 961 CCTTCAGAACTCATTTTACCCTTGATGTTTATGACTTCTACAGATCATCAGCAATGAA 1020  
DB 1150 CCTTCAGAACTCATTTTACCCTTGATGTTTATGACTTCTACAGATCATCAGCAATGAA 1209  
QY 1021 GAAGTTACAGACACATTTGCTATTTCCTCAATATAGATAAAAAAGACTGCACATTTATTACA 1080  
DB 1210 GAAGTTACAGACACATTTGCTATTTCCTCAATATAGATAAAAAAGACTGCACATTTATTACA 1269  
QY 1081 AAAGGCACCTGGGAAGTCACTCGGATAGAAGCTCTAACCCAGTGATTTATCTATCTACTTAC 1140  
DB 1270 AAAGGCACCTGGGAAGTCACTCGGATAGAAGCTCTAACCCAGTGATTTATCTATCTACTTAC 1329  
QY 1141 AGTAATGAATATAAAGGAATGCCAGGAGAGGATCTTTTAAATCCAACTTATTGAC 1200  
DB 1330 AGTAATGAATATAAAGGAATGCCAGGAGAGGATCTTTTAAATCCAACTTATTGAC 1389  
QY 1201 TATACAAAGTGACATGCTCAGTTGTGAGCTGAATCCGGAAGGTGTCAGTACTATTCT 1260  
DB 1390 TATACAAAGTGACATGCTCAGTTGTGAGCTGAATCCGGAAGGTGTCAGTACTATTCT 1449  
QY 1261 GTGTCATTCAGTAAAGAGCGGAAGTATTATCAGCTGAGATGTTCCGGTCTGCGCC 1320  
DB 1450 GTGTCATTCAGTAAAGAGCGGAAGTATTATCAGCTGAGATGTTCCGGTCTGCGCC 1509  
QY 1321 CTCTATCTCTACACAGCAGCGTGAATGATTAAGGGCTGAGAGTCTCTGGAAGACAATTCA 1380  
DB 1510 CTCTATCTCTACACAGCAGCGTGAATGATTAAGGGCTGAGAGTCTCTGGAAGACAATTCA 1569  
QY 1381 GCTTTGGATAAAATGCTCAGAAATGTCAGATGCCCTCCAAAAAACTGGACTTCAATTAT 1440  
DB 1570 GCTTTGGATAAAATGCTCAGAAATGTCAGATGCCCTCCAAAAAACTGGACTTCAATTAT 1629  
QY 1441 TTGAATGAAAACAAAATTTTGGTATCAGATGATCTTGTGCTCTCTCTTGTGATAAATCCAAAG 1500













TPITKTWEVIGIEALTSYLYVINSIEYKMGPRGRNLKYLQISLDYTKVTCLSELNPE  
RCQYVSFSKEAKYQRCSPGLPLVTLHSSVNDKGLRVLEDSALDKMLQNVMP  
SKQDIIILNETKFMQLPDKPKIPULLLVYAGPCSKQADTFRNLNATILYA  
STENIIIVASFDGRSGYQGDKIIMHAINRRLGTFEVEDQIEARQFSKMGFVNDKRIAI  
WMSYGGVYTSNVLGSGSVKPCJAVAPSVRWEYDVSYYTERYMLPTPEDNLDHYR  
NSTVMSRAENFKQVLYLTHGTADDNVHFQSSAQISKALVDVGVDFQAMWYTDDEHGI  
ASSTAHQHIYTHMSHFIKQCFSLP"  
122..1657

misc\_feature

/note="DPP4"

/note="DPP4" N term; Region: Dipeptidyl peptidase IV (DPP  
IV) N-terminal region. This family is an alignment of the  
region to the N-terminal side of the active site"  
/db\_xref="CDD:pfam00930"

ORIGIN

Query Match 99.9%; Score 2182.4; DB 8; Length 2431;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2183; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	AGTCGAAAACCTACACTCTACTGATTTACTTTAAAAAATCTATAGACTGAAGTTATAC	60
DB	110	AGTCGAAAACCTTACACTCTACTGATTTACTTTAAAAAATCTATAGACTGAAGTTATAC	169
QY	61	TCCTTAAAGATGGATTTAGATCATGATATCTCTACAAACAAAGAAAATAATCTTGGTA	120
DB	170	TCCTTAAAGATGGATTTAGATCATGATATCTCTACAAACAAAGAAAATAATCTTGGTA	229
QY	121	TTCAATGCTGAATATGGAAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGATTT	180
DB	230	TTCAATGCTGAATATGGAAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGATTT	289
QY	181	GGACATTTCTCAATGATTTTCAATATCTCTGATGGCAGTTTATCTCTTAGAATAC	240
DB	290	GGACATTTCTCAATGATTTTCAATATCTCTGATGGCAGTTTATCTCTTAGAATAC	349
QY	241	AACTAGCTGAAGCAATGGAGGCAATCTCTACACAGCTTCATATGACATTTATGATTTAAAT	300
DB	350	AACTAGCTGAAGCAATGGAGGCAATCTCTACACAGCTTCATATGACATTTATGATTTAAAT	409
QY	301	AAAAGCAGCTGATTAACAAGAGAGGATTCACAAACACACACAGTGGGTCACTGGTCA	360
DB	410	AAAAGCAGCTGATTAACAAGAGAGGATTCACAAACACACACAGTGGGTCACTGGTCA	469
QY	361	CCAGTGGGTCAATAATGGCATATGTTGGACAAATGACATTTATGTTAAATTCGACCA	420
DB	470	CCAGTGGGTCAATAATGGCATATGTTGGACAAATGACATTTATGTTAAATTCGACCA	529
QY	421	AATTTACCAAGTTTACAGATACATGGACGGGAAAGAGATATATATATATGGAATA	480
DB	530	AATTTACCAAGTTTACAGATACATGGACGGGAAAGAGATATATATATATGGAATA	589
QY	481	ACTGACTGGGTTTATGAAGAGAAAGTCTTCACTGCTACTCTGCTGTGGGTCTCCA	540
DB	590	ACTGACTGGGTTTATGAAGAGAAAGTCTTCACTGCTACTCTGCTGTGGGTCTCCA	649
QY	541	AACGGCACTTTTATGACATATGCCAATTTAACGACACAGAGTCCCACTTTAGATAC	600
DB	650	AACGGCACTTTTATGACATATGCCAATTTAACGACACAGAGTCCCACTTTAGATAC	709
QY	601	TCCTTCTACTCTGATGAGTCACTGACGACCTCCAAAGACCTGACGGTTCATATCCAAAG	660
DB	710	TCCTTCTACTCTGATGAGTCACTGACGACCTCCAAAGACCTGACGGTTCATATCCAAAG	769
QY	661	GCAGGAGCTGTGAATCCAACTGTAAAGTCTTTGTTGTTAAATACAGACTCTCTCAGTCA	720
DB	770	GCAGGAGCTGTGAATCCAACTGTAAAGTCTTTGTTGTTAAATACAGACTCTCTCAGTCA	829
QY	721	GTCAACCAATGCAACTCCATACAAATCACTGCTCTGCTCTCTATGTTGATGGGATCAC	780
DB	830	GTCAACCAATGCAACTCCATACAAATCACTGCTCTGCTCTCTATGTTGATGGGATCAC	889
QY	781	TACTTGTGTGATGTGACATGGGCAACACAAAGAAATTTCTTTGCAAGTGGCTCAGGAGG	840

DB	890	TACTTGTGTGATGTGACATGGGCAACACACAAAGAAATTTCTTTGCAGTGGCTCAGGAGG	949
QY	841	ATTGAGAACTATTCCGCTCATGGATATTTGTGACTATGATGATGATCCAGTGGAGATGGAC	900
DB	950	ATTGAGAACTATTCCGCTCATGGATATTTGTGACTATGATGATGATCCAGTGGAGATGGAC	1009
QY	901	TGCTTAGTGGCAGCGCAACACATTTGAAATGAGTACTACTGGCTGGGTTGGAAAGATTTAGG	960
DB	1010	TGCTTAGTGGCAGCGCAACACATTTGAAATGAGTACTACTGGCTGGGTTGGAGATTTAGG	1069
QY	961	CTTTCAGAACCTCAATTTTACCCTTGTATGATGATGATGATGATGATGATGATGATGATGAT	1020
DB	1070	CTTTCAGAACCTCAATTTTACCCTTGTATGATGATGATGATGATGATGATGATGATGATGAT	1129
QY	1021	GAGGTTACAGACACATTTTGTCTATTTCCAAATAGATTAATAAAGACTGCACATTTATTAACA	1080
DB	1130	GAGGTTACAGACACATTTTGTCTATTTCCAAATAGATTAATAAAGACTGCACATTTATTAACA	1189
QY	1081	AAAGGCACCTCGGAAGTCAATCGGATAGAAGCTCTTAACAGTGAATATCTATCTACTACATT	1140
DB	1190	AAAGGCACCTCGGAAGTCAATCGGATAGAAGCTCTTAACAGTGAATATCTATCTACTACATT	1249
QY	1141	AGTAAATGAATATAAGGAATCCAGGAGGAAGGAATCTTTATAAATCCAACTTTATGAC	1200
DB	1250	AGTAAATGAATATAAGGAATCCAGGAGGAAGGAATCTTTATAAATCCAACTTTATGAC	1309
QY	1201	TATACAAAGTACATGCTCAGTTGTGAGCTGAATCCGGAAGGTGTCAGTACTATTCT	1260
DB	1310	TATACAAAGTACATGCTCAGTTGTGAGCTGAATCCGGAAGGTGTCAGTACTATTCT	1369
QY	1261	GTGTCATTTCAGTAAAGAGCGGAAGTATTTATCAGCTGAGATGTTCCGGTCTCTGCTGCC	1320
DB	1370	GTGTCATTTCAGTAAAGAGCGGAAGTATTTATCAGCTGAGATGTTCCGGTCTCTGCTGCC	1429
QY	1321	CTCTATCTCTACACAGCAGCGTGAATGATATAAGGGCTGAGAGTCTCTGGAAGACATTTCA	1380
DB	1430	CTCTATCTCTACACAGCAGCGTGAATGATATAAGGGCTGAGAGTCTCTGGAAGACATTTCA	1489
QY	1381	GCTTTGGATAAATGCTGACAGATGCTCAGATGCCCTCCAAAACACTGGACTTCATTTAT	1440
DB	1490	GCTTTGGATAAATGCTGACAGATGCTCAGATGCCCTCCAAAACACTGGACTTCATTTAT	1549
QY	1441	TTGAATGAAAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGTATAAATCCAAG	1500
DB	1550	TTGAATGAAAACAAAATTTTGGTATCAGATGATCTTGGCTCTCTCATTTTGTATAAATCCAAG	1609
QY	1501	AAATATCTCTACTATTAAGTGTGTATGCAAGGCCCATGTAGTCAAAAAGCAGACACTGTC	1560
DB	1610	AAATATCTCTACTATTAAGTGTGTATGCAAGGCCCATGTAGTCAAAAAGCAGACACTGTC	1669
QY	1561	TTGAGCTGAATCTGGGCCACTTACCTTCCAGCACAGAAACATTTATAGTACTAGCTTT	1620
DB	1670	TTGAGCTGAATCTGGGCCACTTACCTTCCAGCACAGAAACATTTATAGTACTAGCTTT	1729
QY	1621	GATGCGCAGAGAAAGTGGTTTACCAAGGAGATTAAGATCATGCAATGCAATCAACAGAGACTG	1680
DB	1730	GATGCGCAGAGAAAGTGGTTTACCAAGGAGATTAAGATCATGCAATGCAATCAACAGAGACTG	1789
QY	1681	GGAAATTTGAAGTTGAAGATCAAAATTTGAAGCAGCAGACAAATTTTCAAAAATGGGATTT	1740
DB	1790	GGAAATTTGAAGTTGAAGATCAAAATTTGAAGCAGCAGACAAATTTTCAAAAATGGGATTT	1849
QY	1741	GTGGCAACAAACGAAATTTGGGGCTGGTTCATATGGAGGTGATGATGATGATGATGATGATG	1800
DB	1850	GTGGCAACAAACGAAATTTGGGGCTGGTTCATATGGAGGTGATGATGATGATGATGATGATG	1909
QY	1801	GTCTTGGGATCGGGAAGTGGGCTGTTCAAGTGTGGAATAGCGCTGGCGCTCTATCCCCGG	1860
DB	1910	GTCTTGGGATCGGGAAGTGGGCTGTTCAAGTGTGGAATAGCGCTGGCGCTCTATCCCCGG	1969
QY	1861	TGGGAGTACTATGACTCAGTGTACAGAAAGTTCATGAGTGGGTCTCTCCAACTCCAGAGAC	1920
DB	1970	TGGGAGTACTATGACTCAGTGTACAGAAAGTTCATGAGTGGGTCTCTCCAACTCCAGAGAC	2029



[illegible]



Db 1571 GCTTTGGATAAAATGCTGCAGAAATGTCCAGATGCCCTCCAAAAAATCGGACTTCATTATT 1630  
QY 1441 TTGAATGAAACAAAATTTTGGTATCAGATGATCTTGCTCCTCATTTTGATATAATCCAAAG 1500  
Db 1631 TTGAATGAAACAAAATTTTGGTATCAGATGATCTTGCTCCTCATTTTGATATAATCCAAAG 1690  
QY 1501 AAATATCCTCTACTATTAGATGTGTATGCAGGCCCATGTAGTCAAAAAGCAGACACTGTC 1560  
Db 1691 AAATATCCTCTACTATTAGATGTGTATGCAGGCCCATGTAGTCAAAAAGCAGACACTGTC 1750  
QY 1561 TTCAGACTGAACTGGGCCACTTACCTTGCAGCACAAGAAACATTTATAGTAGCTAGCTTT 1620  
Db 1751 TTCAGACTGAACTGGGCCACTTACCTTGCAGCACAAGAAACATTTATAGTAGCTAGCTTT 1810  
QY 1621 GATGGCAGGAAGTGGTTACCAAGGAGATAGATCATGTCATCAACAGAGAGACTG 1680  
Db 1811 GATGGCAGGAAGTGGTTACCAAGGAGATAGATCATGTCATCAACAGAGAGACTG 1870  
QY 1681 GGAACAATTTTGAAGTTGAAGATCAAAATTTGAAGCAGCCAGACAATTTTCAAAAATGGGATTT 1740  
Db 1871 GGAACAATTTTGAAGTTGAAGATCAAAATTTGAAGCAGCCAGACAATTTTCAAAAATGGGATTT 1930  
QY 1741 GTGGACAAACAAACGAATTCGAATTTGGGCTGGTCAATATGGAGGGTACGTAACTCAATG 1800  
Db 1931 GTGGACAAACAAACGAATTCGAATTTGGGCTGGTCAATATGGAGGGTACGTAACTCAATG 1990  
QY 1801 GTCTGGGATCGGGAAGTGGCTGTTCAGATGTGGAATAGCCGTGGCGCTGTATCCCGG 1860  
Db 1991 GTCTGGGATCGGGAAGTGGCTGTTCAGATGTGGAATAGCCGTGGCGCTGTATCCCGG 2050  
QY 1861 TGGGAGTACTATGACTCAGTGTACACAGAACGTTTACATGGGTCTCCCAACTCCAGAAGAC 1920  
Db 2051 TGGGAGTACTATGACTCAGTGTACACAGAACGTTTACATGGGTCTCCCAACTCCAGAAGAC 2110  
QY 1921 AACCTTGACCATTTACAGAAATTTCAACAGTCATGAGCAGAGCTGAAAAATTTTAAACAAGTT 1980  
Db 2111 AACCTTGACCATTTACAGAAATTTCAACAGTCATGAGCAGAGCTGAAAAATTTTAAACAAGTT 2170  
QY 1981 GAGTACCTCTTATTCATGGAAACAGCAGATGATTAACGTTTCACTTTCAGCAGTCAGCTCAG 2040  
Db 2171 GAGTACCTCTTATTCATGGAAACAGCAGATGATTAACGTTTCACTTTCAGCAGTCAGCTCAG 2230  
QY 2041 ATCTCCAAAGCCCTGGTCATGTTGGAGTGGATTTCCAGGCAATGTGGTATATCTGATGAA 2100  
Db 2231 ATCTCCAAAGCCCTGGTCATGTTGGAGTGGATTTCCAGGCAATGTGGTATATCTGATGAA 2290  
QY 2101 GACCATGGAATAGCTAGCAGCAGCAGCACCACATATATATACCCACATGAGCCACTTC 2160  
Db 2291 GACCATGGAATAGCTAGCAGCAGCAGCACCACATATATATATACCCACATGAGCCACTTC 2350  
QY 2161 ATAAACAATGTTTCTTTACCT 2184  
Db 2351 ATAAACAATGTTTCTTTACCT 2374

RESULT 9  
CQ776356  
LOCUS CQ776356 3445 bp DNA linear PAT 11-MAR-2004  
DEFINITION Sequence 42 from Patent EP1394274.  
ACCESSION CQ776356  
VERSION CQ776356.1 GI:45379746  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1  
Ohtani,N., Sugita,Y., Yamaya,M., Kubo,H., Nagai,H. and Izuohara,K.  
AUTHORS Methods of testing for bronchial asthma or chronic obstructive  
TITLE pulmonary disease  
JOURNAL Patent: EP 1394274-A 42 03-MAR-2004;  
Genox Research, Inc. (JP)

FEATURES  
source Location/Qualifiers  
1. 3445  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 99.9%; Score 2182.4; DB 6; Length 3445;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2183; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 AGTCGCAAAACCTTACACTCTAACTGATTACTTTAAAAAATACTTATAGACTGAAGTTATAC 60  
Db 225 AGTCGCAAAACCTTACACTCTAACTGATTACTTTAAAAAATACTTATAGACTGAAGTTATAC 284  
QY 61 TCCTTAAAGATGGATTCAGATCATGAATATCTCTACAAACAGAAATAATATCTTGGTA 120  
Db 285 TCCTTAAAGATGGATTCAGATCATGAATATCTCTACAAACAGAAATAATATCTTGGTA 344  
QY 121 TTCATGCTGAATATGGAAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 180  
Db 345 TTCATGCTGAATATGGAAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 404  
QY 181 GGACATTTCTCAATGATTATTCAATATCTCTGATGGCAGTTTATTCTCTTAGAATAC 240  
Db 405 GGACATTTCTCAATGATTATTCAATATCTCTGATGGCAGTTTATTCTCTTAGAATAC 464  
QY 241 AACTAGTGAAGCAATGGAGGCAATTCACACAGCTTCATATGACATTTATGATTTAAAT 300  
Db 465 AACTAGTGAAGCAATGGAGGCAATTCACACAGCTTCATATGACATTTATGATTTAAAT 524  
QY 301 AAAAGCGCAGCTGATTACAGAGAGAGGATTCCAAACAAACACACAGTGGGTACATGGTCA 360  
Db 525 AAAAGCGCAGCTGATTACAGAGAGAGGATTCCAAACAAACACACAGTGGGTACATGGTCA 584  
QY 361 CCAGTGGGTCTATAAATTTGGCATATGTTTGGAAACAATGACATTTATGTTAAAAATGAAACA 420  
Db 595 CCAGTGGGTCTATAAATTTGGCATATGTTTGGAAACAATGACATTTATGTTAAAAATGAAACA 644  
QY 421 AATTACCAGTGTACAGATTCACATGGAGCGGGAAGAAGATATATAATATTAATGGAATA 480  
Db 645 AATTACCAGTGTACAGATTCACATGGAGCGGGAAGAAGATATATAATATTAATGGAATA 704  
QY 481 ACTGACTGGTTTATGAAGAGAGAGTCTTCAGTGCCTACTCTGCTCTGTGGTGGTCTCCA 540  
Db 705 ACTGACTGGTTTATGAAGAGAGAGTCTTCAGTGCCTACTCTGCTCTGTGGTGGTCTCCA 764  
QY 541 AACGGCAGCTTTTATAGCATATGCCCAATTTAAACGACACAGAAAGTCCCACCTTATGGAATAC 600  
Db 765 AACGGCAGCTTTTATAGCATATGCCCAATTTAAACGACACAGAAAGTCCCACCTTATGGAATAC 824  
QY 601 TCCTTCTACTCTGATGAGTCACTGACGTACCCAAAGACTGTACGGGTCCCATATCCAAAG 660  
Db 825 TCCTTCTACTCTGATGAGTCACTGACGTACCCAAAGACTGTACGGGTCCCATATCCAAAG 884  
QY 661 GCAGGAGCTGTGAATCCAACTGTAAGTTCTTTGTTGTAATAACAGACTCTCTCAGCTCA 720  
Db 885 GCAGGAGCTGTGAATCCAACTGTAAGTTCTTTGTTGTAATAACAGACTCTCTCAGCTCA 944  
QY 721 GTCAACCAATGCAACTTCCATACAAATCACTGCTCCTGCTTCTATGTTGATAGGGGATCAC 780  
Db 945 GTCAACCAATGCAACTTCCATACAAATCACTGCTCCTGCTTCTATGTTGATAGGGGATCAC 1004  
QY 781 TACTTGTGTGATGTGACATGGGCAACACAGAAAGAAATTTCTTTCGAGTGGCTCAGGAGG 840  
Db 1005 TACTTGTGTGATGTGACATGGGCAACACAGAAAGAAATTTCTTTCGAGTGGCTCAGGAGG 1064  
QY 841 ATTCAAGAACTATTCCGTTCATGGATATTTGTGACTATGATGAATCCAGTGGGAAGATGGAAC 900  
Db 1065 ATTCAAGAACTATTCCGTTCATGGATATTTGTGACTATGATGAATCCAGTGGGAAGATGGAAC 1124  
QY 901 TGCTTTAGTGGCACGGCAACACATTTGAAATGAGTACTACTGGCTGGGTTCGGAAGATTTAGG 960



Db	645	 AATTTACAAAGTTACAGAAATCACATGCAGCGGGAAAGAGATATAATATATAATGAATA	704
Qy	481	 ACTGACCTGGGTTTATGAAGAGGAAGCTCTTCAGTGCCTACTCTGCTCTGCTGGTCTCCCA	540
Db	705	 ACTGACCTGGGTTTATGAAGAGGAAGTCTTCAGTGCCTACTCTGCTCTGCTGGTCTCCCA	764
Qy	541	 AACGGCACTTTTATGACATATGCCCAATTTAAACGACACAGAAAGTCCCACTTATTGAATAC	600
Db	765	 AACGGCACTTTTATGACATATGCCCAATTTAAACGACACAGAAAGTCCCACTTATTGAATAC	824
Qy	601	 TCCTTCTACTCTGATGAGTCACTGCAGTACCCAAAGA CTGTACGGGTTCCATATPCAAAG	660
Db	825	 TCCTTCTACTCTGATGAGTCACTGCAGTACCCAAAGACTGTACGGGTTCCATATPCAAAG	884
Qy	661	 GCAGGAGCTGTAATCCAACTGTAAAGTCTTTTGTGTAAATACAGACTCTCTCAGTCA	720
Db	885	 GCAGGAGCTGTGAATCCAACTGTAAAGTCTTTTGTGTAAATACAGACTCTCTCAGTCA	944
Qy	721	 GTCAACCAATGCAACTTCCATACAATCACTGCTCCTCTCTATGTTGTAGTAGGGATCAC	780
Db	945	 GTCAACCAATGCAACTTCCATACAATCACTGCTCCTCTCTATGTTGTAGTAGGGATCAC	1004
Qy	781	 TACTTGTGTGATGTGACATGGGCAACA CAAGAAAGAAATTTCTTTGCACTGGCTCAGGAGG	840
Db	1005	 TACTTGTGTGATGTGACATGGGCAACA CAAGAAAGAAATTTCTTTGCACTGGCTCAGGAGG	1064
Qy	841	 ATTCAGAACTATTCCGGTCATGGATTTTGTGACTATGATGAATCCAGTGGAGATGGAAC	900
Db	1065	 ATTCAGAACTATTCCGGTCATGGATTTTGTGACTATGATGAATCCAGTGGAGATGGAAC	1124
Qy	901	 TGCTTAGTGGCAGCGCAACA CATTGAAATCAGTACTACTGGCTGGGTTGGAAGATTTAGG	960
Db	1125	 TGCTTAGTGGCAGCGCAACA CATTGAAATCAGTACTACTGGCTGGGTTGGAAGATTTAGG	1184
Qy	961	 CCTTCAGAACCTCATTTTACCCTTGTGATGGTAATAGTCTTACAAGATCATCAGCAATGAA	1020
Db	1185	 CCTTCAGAACCTCATTTTACCCTTGTGATGGTAATAGTCTTACAAGATCATCAGCAATGAA	1244
Qy	1021	 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATATAAANAAGACTGCACATTTATTACA	1080
Db	1245	 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATATAAANAAGACTGCACATTTATTACA	1304
Qy	1081	 AAAGGCACCTGGGAAGTCACTCGGATAGAAGCTCTTAACCAAGTGAATATCTATCTACATTT	1140
Db	1305	 AAAGGCACCTGGGAAGTCACTCGGATAGAAGCTCTTAACCAAGTGAATATCTATCTACATTT	1364
Qy	1141	 AGTAATGAATATAAAGGAATGCCAGGAGGAAGAAATCTTTTATAAAATCCAACTTATTGAC	1200
Db	1365	 AGTAATGAATATAAAGGAATGCCAGGAGGAAGAAATCTTTTATAAAATCCAACTTATTGAC	1424
Qy	1201	 TATCAAAAGTGACATGCCCTCAGTTGTGAGCTGAATCCGGAAGGTTGTCAGTACTATTCT	1260
Db	1425	 TATCAAAAGTGACATGCCCTCAGTTGTGAGCTGAATCCGGAAGGTTGTCAGTACTATTCT	1484
Qy	1261	 GTGTCACTTCAGTAAAGAGGGCAAGTATTATCAGCTGAGATGTTCCGGTCTCTGGTCTGCC	1320
Db	1485	 GTGTCACTTCAGTAAAGAGGGCAAGTATTATCAGCTGAGATGTTCCGGTCTCTGGTCTGCC	1544
Qy	1321	 CTCTATACTCTACACAGCAGCGTGAATGATAAAGGGCTGAGAGTCTCTGGGAAGACAATTC	1380
Db	1545	 CTCTATACTCTACACAGCAGCGTGAATGATAAAGGGCTGAGAGTCTCTGGGAAGACAATTC	1604
Qy	1381	 GCTTTGATATAATGCTGCAAGATGTCAGATGCCCTCCAAAACCTGGACCTCATTTATT	1440
Db	1605	 GCTTTGATATAATGCTGCAAGATGTCAGATGCCCTCCAAAACCTGGACCTCATTTATT	1664
Qy	1441	 TTGAATGAAACAAAATTTTGTATCAGATGATCTTGCTCCTCTCATTTTGTATAAATCCAAG	1500
Db	1665	 TTGAATGAAACAAAATTTTGTATCAGATGATCTTGCTCCTCTCATTTTGTATAAATCCAAG	1724
Qy	1501	 AAATPATCCTCTACTATTAGATGTGTATGCAAGGCCCATGTGATCAAAAGCAGACACTGTC	1560

[illegible]

[illegible]





[illegible]



Qy	1201	TATACAAAAGTGACATGCCTCAGTTGTGAGCTGAATCCGAAAAGGTGTGAGTACTATTCT	1260	
Db	1425	TATACAAAAGTGACATGCCTCAGTTGTGAGCTGAATCCGAAAAGGTGTGAGTACTATTCT	1484	
Qy	1261	GTGTCAATTCAGTAAAGAGGCGAAGTATTATCAGCTGAGATGTTCCGGTCTCTGTCTGCC	1320	
Db	1485	GTGTCAATTCAGTAAAGAGGCGAAGTATTATCAGCTGAGATGTTCCGGTCTCTGTCTGCC	1544	
Qy	1321	CTCTATACTCTACACAGCAGCGTGAATGATAAAGGGCTGAGAGTCTCTGGAAGCAATTC	1380	
Db	1545	CTCTATACTCTACACAGCAGCGTGAATGATAAAGGGCTGAGAGTCTCTGGAAGCAATTC	1604	
Qy	1381	GCTTTGGATAAATGTCTGCAGAAATGTCCAGATGCCCTCCAAAAAATCGACTTCATTATT	1440	
Db	1605	GCTTTGGATAAATGTCTGCAGAAATGTCCAGATGCCCTCCAAAAAATCGACTTCATTATT	1664	
Qy	1441	TTGAATGAAACAAAATTTTGGTATCAGATGATCTTGCTCTCTCATTTTGTGATAAATCCAAG	1500	
Db	1665	TTGAATGAAACAAAATTTTGGTATCAGATGATCTTGCTCTCTCATTTTGTGATAAATCCAAG	1724	
Qy	1501	AAATATCCTCTCATATTAGATGTGTATGACGGCCCATGTAGTCAAAAAGCAGACACTGTC	1560	
Db	1725	AAATATCCTCTCATATTAGATGTGTATGACGGCCCATGTAGTCAAAAAGCAGACACTGTC	1784	
Qy	1561	TTCAGACTGAATCGGGCCACTTACCTTGCAGACACAGAAAAACATTATAGTAGCTAGCTTT	1620	
Db	1785	TTCAGACTGAATCGGGCCACTTACCTTGCAGACACAGAAAAACATTATAGTAGCTAGCTTT	1844	
Qy	1621	GATGGCAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGCAATCAACAGAAAGACTG	1680	
Db	1845	GATGGCAGAGGAAGTGGTTTACCAAGGAGATAAGATCATGCAATCAACAGAAAGACTG	1904	
Qy	1681	GGAACTTTGCAAGTTGAAGATCAAAATTGAAGCAGCCAGACAATTTTCAAAAATGGGATTT	1740	
Db	1905	GGAACTTTGCAAGTTGAAGATCAAAATTGAAGCAGCCAGACAATTTTCAAAAATGGGATTT	1964	
Qy	1741	GTGGACAAACAAACGAATTTGCAATTTGGGGCTGGTCATATGGAGGGTACGTAACTCAATG	1800	
Db	1965	GTGGACAAACAAACGAATTTGCAATTTGGGGCTGGTCATATGGAGGGTACGTAACTCAATG	2024	
Qy	1801	GTCTTGGGATCGGGAAGTGGCGTGTTCAAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG	1860	
Db	2025	GTCTTGGGATCGGGAAGTGGCGTGTTCAAAGTGTGGAATAGCCGTGGCGCTGTATCCCGG	2084	
Qy	1861	TGGGAGTACTATGACTCAGTGTACAGAACGTTACATGGGTCTCCCAACTCCAGAAGAC	1920	
Db	2085	TGGGAGTACTATGACTCAGTGTACAGAACGTTACATGGGTCTCCCAACTCCAGAAGAC	2144	
Qy	1921	AACCTTGACCAATTACAGAAATTTCAACAGTCATGAGCAGAGCTGAAAATTTTAAACAAGTT	1980	
Db	2145	AACCTTGACCAATTACAGAAATTTCAACAGTCATGAGCAGAGCTGAAAATTTTAAACAAGTT	2204	
Qy	1981	GAGTACCTCTCTTATTCATGGAACAGCAGATGATAACGTTTCACTTTTCCAGCAGTCAG	2040	
Db	2205	GAGTACCTCTCTTATTCATGGAACAGCAGATGATAACGTTTCACTTTTCCAGCAGTCAG	2264	
Qy	2041	ATCTCCAAAGCCCTGGTCGATGTTGGAGTGGATTTCCAGGCAATGTGGTATATCTGATGAA	2100	
Db	2265	ATCTCCAAAGCCCTGGTCGATGTTGGAGTGGATTTCCAGGCAATGTGGTATATCTGATGAA	2324	
Qy	2101	GACCATGGATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	2160	
Db	2325	GACCATGGATAGCTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	2384	
Qy	2161	ATAAAAACAAATGTTTCTCTTTACCT 2184		
Db	2385	ATAAAAACAAATGTTTCTCTTTACCT 2408		
RESULT 14				
CS041831				
LOCUS				
CS041831 3445 bp DNA linear PAT 22-MAR-2005				

DEFINITION	Sequence 2385 from Patent WO2005019258.
ACCESSION	CS041831
VERSION	CS041831.1
KEYWORDS	GI:61849188
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Abbas, A., Clark, H., Ouyang, W., Williams, P. M., Wood, W. I. and Wu, T. D. Compositions and methods for the treatment of immune related diseases
JOURNAL	Patent: WO 2005019258-A 2385 03-MAR-2005;
FEATURES	Genentech, Inc. (US)
source	Location/Qualifiers
	1..3445
	/organism="Homo sapiens"
	/mol_type="unassigned DNA"
	/db_xref="taxon:9606"
ORIGIN	
	Query Match 99.9%; Score 2182.4; DB 6; Length 3445;
	Best Local Similarity 99.9%; Pred. No. 0;
	Matches 2183; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	1 AGTCGCAAACTTACACTCTAACTGATTACTTAAAAAATCTTATAGACTGAAGTTATAC 60
Db	225 AGTCGCAAACTTACACTCTAACTGATTACTTAAAAAATCTTATAGACTGAAGTTATAC 284
QY	61 TCCTTAAAGATGGATTTTTCAGATCATGAATATCTCTACAAACAAAGAAAATAATCTTGGTA 120
Db	285 TCCTTAAAGATGGATTTTTCAGATCATGAATATCTCTACAAACAAAGAAAATAATCTTGGTA 344
QY	121 TTCATGCTGAATATGGAAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 180
Db	345 TTCATGCTGAATATGGAAACAGCTCAGTTTCTTGGAGAACAGTACATTTGATGAGTTT 404
QY	181 GGACATTTCTATCAATGATTATTCATATCTCTGATGGCAGTTTATCTCTTAGAATAC 240
Db	405 GGACATTTCTATCAATGATTATTCATATCTCTGATGGCAGTTTATCTCTTAGAATAC 464
QY	241 AACTACGTGAAGCAATGAGGCAATTCCTACACAGCTTCATATGACATTTATGATTTAAAT 300
Db	465 AACTACGTGAAGCAATGAGGCAATTCCTACACAGCTTCATATGACATTTATGATTTAAAT 524
QY	301 AAAAGGACGTGATTTACAGAGAGAGAGGATTCACAAACACACACAGTGGGTCACTGGTCA 360
Db	525 AAAAGGACGTGATTTACAGAGAGAGAGGATTCACAAACACACACAGTGGGTCACTGGTCA 584
QY	361 CCAGTGGGTCAATAATTTGGCATATGTTTGGACANTGACATTTATGTTAAAAATTCACCA 420
Db	585 CCAGTGGGTCAATAATTTGGCATATGTTTGGACANTGACATTTATGTTAAAAATTCACCA 644
QY	421 AATTTACCAGTTTACAGAAATCACTGGACGGGAAAGAGATATATATATATGAAATA 480
Db	645 AATTTACCAGTTTACAGAAATCACTGGACGGGAAAGAGATATATATATGAAATA 704
QY	481 ACTGACTGGGTTTATGAAGAGAGAGTCTTCAGTGCCTACTCTGCTCTGTGGGTCTCCA 540
Db	705 ACTGACTGGGTTTATGAAGAGAGAGTCTTCAGTGCCTACTCTGCTCTGTGGGTCTCCA 764
QY	541 AACGGCACTTTTATGACATATGCCCAATTTTAAACGACACAGAGTCCCACCTTATGAAATAC 600
Db	765 AACGGCACTTTTATGACATATGCCCAATTTTAAACGACACAGAGTCCCACCTTATGAAATAC 824
QY	601 TCCTTCTACTCTGATGAGTCACTGAGTACCCAAAGACTGTACGGGTCCATATCCAAAG 660
Db	825 TCCTTCTACTCTGATGAGTCACTGAGTACCCAAAGACTGTACGGGTCCATATCCAAAG 884
QY	661 GCAGGAGCTGGAATCCAACTGTAAGTTCTTTGTTGTAATAATACAGACTCTCTCAGCTCA 720
Db	885 GCAGGAGCTGTAAGTTCTTTGTTGTAATAATACAGACTCTCTCAGCTCA 944



NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: [nisc.mgc@nih.gov](mailto:nisc.mgc@nih.gov)

Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,  
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,  
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaepi, R.,  
Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,  
McDowell, J., Pearson, R., Stantzirop, S., Thomas, P.J., Touchman, J.W.,  
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,  
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK Plate: 140 Row: 1 Column: 9  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 18765693.

# FEATURES

source

1. .3913

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="MGC:74748 IMAGE:6156880"

/tissue\_type="uterus, leiomyosarcoma"

/clone\_lib="NIH MGC\_71"

/lab\_host="DH10B"

/note="Vector: pCMV-SPORT6"

1. .3913

/gene="DPP4"

/note="synonyms: ADP2, CD26, DPPIV, ADABP, TP103"

/db\_xref="GeneID:1803"

/db\_xref="MIM:102720"

562. .2862

/gene="DPP4"

/codon\_start=1

/product="dipeptidylpeptidase IV"

/protein\_id="AAH65265.1"

/db\_xref="GI:40850969"

/db\_xref="GeneID:1803"

/db\_xref="MIM:102720"

/translation="MKTWKVLLGLLGAALVTIITVPVLLNKGTDADTSRKYIT  
LTDYLNTRYRLKLSLRWISDHEYLKQENILVFNAYGNSVPLENSTDFGHSI  
NDYISPOGQFILLENYVYKQWRHSYTA5DIYDLNKRQLITEERIPNNQWTVSPV  
GHKJAYVNMNDIYVIEPLNPSYRITWTGKEDIYNGIDTWYEEVFSAYSALNWSF  
SGTFLAYQFNDTEVPLIEBYSFSDLESQYKPTVRVPYKAGAVNPYKFFVYNTDSL  
NSVTWATSIOITAPASMLIGDHYLGDVTWATQERISLOWLRRIQNYVMDICDYDESS  
GRNCLVARQHIEMSTGWGVRFRSEPHFTLDGNSFYKIIISNEBGRHICVFOIDKK  
DCTFYTKGTWIGVIGEALTSYLYVYISNEYKMGPRGNLYKLIQSDYTKVTCLSCELN  
PERQYTSVFSKEAKYIQLRSGGPELPITLHSSVNDKRLVSDNLSALDKRLQNVQ  
MPSKKDLFIILNETFWYQMLPHFDDSKKYPILLDVGPCSKADTVFRLNWTY  
LASTENIIVASFDRSGSGYQGDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRI  
ATGWSYGGYVTSVMVSGSGYFKGCIAPVPSRWBYVDSVYTERYMGILPTPEDNLH  
YNSVTMSRAENFKQVEYLLIHGTADNVHFQSSAQISKALVDVGVDFOAMWYTDH  
GIASSTAHQHIYTHMSHFIKQCFSLP"

## ORIGIN

Query Match 99.9%; Score 2182.4; DB 8; Length 3913;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2183; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
1 AGTCGAAAACCTTACACTCTAAGTATCTTAAAAAATCTTATAGACTGAAGTTATAC 60  
676 AGTCGAAAACCTTACACTCTAAGTATCTTAAAAAATCTTATAGACTGAAGTTATAC 735  
91 TCCTTAAGATGGATTTTCAGATCATGAATATCTCTACAAACAGAAAATAATATCTTGGTA 120

736 TCCTTAAGATGGATTTTCAGATCATGAATATCTCTACAAACAGAAAATAATATCTTGGTA 795  
121 TTCAATGCTGAATATGGAACACAGCTCAGTTCTTGGAGAACAGTACATTTGATGAGTTT 180  
796 TTCAATGCTGAATATGGAACACAGCTCAGTTCTTGGAGAACAGTACATTTGATGAGTTT 855  
181 GGACATTTCTATCAATGATTTATCAATATCTCTGATGGCAGTTATTTCTTTAGAAATAC 240  
856 GGACATTTCTATCAATGATTTATCAATATCTCTGATGGCAGTTATTTCTTTAGAAATAC 915  
241 AACTACGTTGAAGCAATGAGGCAATCTCTACACAGCTTTCATATGACATTTATGATTAAAT 300  
916 AACTACGTTGAAGCAATGAGGCAATCTCTACACAGCTTTCATATGACATTTATGATTAAAT 975  
301 AAAAGGACGCTGATTTACAGAGAGAGGATTTCCAAAACACACACAGTGGGTACATGGTCA 360  
976 AAAAGGACGCTGATTTACAGAGAGAGGATTTCCAAAACACACACAGTGGGTACATGGTCA 1035  
361 CCAGTGGGTCTATAAATTCGCATATCTTTGGAAACATGACATTTATGTTAAAAATTGAACCA 420  
1036 CCAGTGGGTCTATAAATTCGCATATCTTTGGAAACATGACATTTATGTTAAAAATTGAACCA 1095  
421 AATTTACCAAGTTACAGAAATCACATGGACGGGGAAGAGATATATATATATATGAAATA 480  
1096 AATTTACCAAGTTACAGAAATCACATGGACGGGGAAGAGATATATATATATGAAATA 1155  
481 ACTGATCGGTTTATGAAGAGAGAGTCTTTCAGTGGCTACTCTGCTCTGTGGTGGTCTCCA 540  
1156 ACTGATCGGTTTATGAAGAGAGAGTCTTTCAGTGGCTACTCTGCTCTGTGGTGGTCTCCA 1215  
541 AACGGCACTTTTATAGCATATGCCCAATTTTAAACACACAGAGAGTCCCACCTTATGGAATAC 600  
1216 AACGGCACTTTTATAGCATATGCCCAATTTTAAACACACAGAGAGTCCCACCTTATGGAATAC 1275  
601 TCCTTCTACTCTGATGAGTCACTGACAGTACCCAAAGAGCTGTACGGGTTCATATCCAAAG 660  
1276 TCCTTCTACTCTGATGAGTCACTGACAGTACCCAAAGAGCTGTACGGGTTCATATCCAAAG 1335  
661 GCAGAGCTGTGAATCCAACTGTAAAGTCTTTTGTGTTGTAATACAGACTCTCTCAGCTCA 720  
1336 GCAGAGCTGTGAATCCAACTGTAAAGTCTTTTGTGTTGTAATACAGACTCTCTCAGCTCA 1395  
721 GTCCCAATGCAACTTCCATCAAAATCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780  
1396 GTCCCAATGCAACTTCCATCAAAATCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1455  
781 TACTTGTGTGATGTGACATGGGCAACACAGAAAGAAATTTCTTTCAGTGGGTTCAGGAGG 840  
1456 TACTTGTGTGATGTGACATGGGCAACACAGAAAGAAATTTCTTTCAGTGGGTTCAGGAGG 1515  
841 ATTGAGAACTATTCGGTCAATGATATTTGTGATATGATGAATCCAGTGGGAAGATGGAAC 900  
1516 ATTGAGAACTATTCGGTCAATGATATTTGTGATATGATGAATCCAGTGGGAAGATGGAAC 1575  
901 TGCTTAGTGGCACGGCAACACATTCGAAATGAGTACTCTGCTGGTGGTTCGGAAGATTAGG 960  
1576 TGCTTAGTGGCACGGCAACACATTCGAAATGAGTACTCTGCTGGTGGTTCGGAAGATTAGG 1635  
961 CCTTCAGAACTCTCATTTTACCCTTGCATGATAGTCTTCTACAAGATCATCAGCAATGAA 1020  
1636 CCTTCAGAACTCTCATTTTACCCTTGCATGATAGTCTTCTACAAGATCATCAGCAATGAA 1695  
1021 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATAAAAAAGACTGCACATTTATTACA 1080  
1696 GAAGGTTACAGACACATTTGCTATTTCCAAATAGATAAAAAAGACTGCACATTTATTACA 1755  
1081 AAAGGCACTGGGAAGTCTATCGGATAGAGCTCTTAAACAGTATTATCTATCTACTACATT 1140  
1756 AAAGGCACTGGGAAGTCTATCGGATAGAGCTCTTAAACAGTATTATCTATCTACTACATT 1815  
1141 AGTAATGAATATPAAAGGAATGCCAGGAGGAAGAAATCTTTATATAAATCCAACTTATTGAC 1200

Search completed: February 17, 2006, 07:42:14  
Job time : 11148 secs

THIS PAGE BLANK (user)